Author:
Rome, Nick A

Title:
Local-global problems in diophantine geometry

Local-Global Problems in Diophantine Geometry

Nick Rome

April 2020

A dissertation submitted to the University of Bristol in accordance with the requirements for award of the degree of Doctor of Philosophy in the Faculty of Science, School of Mathematics.

Abstract

This thesis is dedicated to the study of rational solutions to Diophantine equations, or equivalently rational points on varieties. In particular, we consider families of equations and ask how frequently rational solutions exist to equations in these families. We also study the frequency of examples of local-global principles, such as the Hasse principle and weak approximation, for equations in these families. We develop an asymptotic formula, in Chapter 3, for the number of biquadratic fields K of bounded discriminant for which the equation $N_{K/\mathbb{Q}}(x) = c$ fails to satisfy the Hasse principle for each $c \in \mathbb{Q}^*$. In Chapter 4, we study the frequency of Hasse principle failures in a particular family of rational surfaces. Families of varieties whose base is a hypersurface of low degree is the subject of Chapter 5. In Chapter 6, we describe the frequency with which conics of the form $aX^2 + bY^2 + cZ^2 = 0$ have non-trivial rational points. Finally, in the last chapter we study weak approximation for certain quadric surface bundles over $\mathbb{P}^2$.

Acknowledgements

That this thesis exists at all is a testament to the patience and generosity (both of time and of insight) of my advisor, Tim Browning. It is an old joke that a PhD is just a very slow way for your advisor to talk to himself, it has been a pleasure to be a part of that conversation.

I am deeply grateful to my friends in the Bristol maths department who made Bristol a home for me and to the members of the Browning group (2019-2020) who made Vienna a new one. To the colleagues, too numerous to name, who have helped me feel welcomed into the wider community of rational points I am indebted.

For useful comments, conversations and complaints on some of the individual projects which make up this thesis I thank Regis de la Bretèche, Jean-Louis Colliot-Thélène, Damaris Schindler, Rachel Newton, Andy Booker and Trevor Wooley.

In Adelina Mânzăteanu, I could not have asked for a better collaborator, conference companion and friend. All that I know about the Brauer–Manin obstruction I owe to the patient tutelage of Julian Lyczak. I am especially grateful to Efthymios Sofos for inviting me to work with him in MPIM, his constant (not always useful) advice and for his unique perspective on mathematics and the world.

My thanks to the EPSRC for funding this PhD, and to the University of Bristol and IST Austria for their support.

Finally, to my family, Tom, Lydia, Mum and Dad, thank you for your support not just in the last 4 years but always. And thank you for never pressing me too hard to explain what it is I actually do.

Author's declaration

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's Regulations and Code of Practice for Research Degree Programmes and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the dissertation are those of the author.

Nick Rome 20/04/2020

Contents

Chapter 1

Introduction

The quest to find rational solutions to polynomial equations is one of the oldest in mathematics if not all of human thought. Figure 1.1 depicts the Babylonian tablet known as Plimpton 322 enumerating integer solutions to the Pythagorean equation $x^2 + y^2 = z^2$ dated as over 3800 years old [97]. To this day there is still great interest in understanding when a polynomial equation with integer coefficients (henceforth referred to as Diophantine) has integral or rational solutions. There is a natural trichotomy of possible answers: either there are no solutions, solutions exist but only finitely many or infinitely many solutions exist. When given a set of equations it is not obvious whether or not to expect any solutions to exist. Indeed in 1901, Hilbert posed this question as the tenth in his celebrated list of the 23 most pressing questions that were eluding the mathematical community. In 1970, this question was resolved for integral solutions by Matiyasevich[86]. Namely, he showed that there cannot exist any algorithm which when given a system of Diophantine equations could decide whether or not there exist integral solutions. The analogous question remains open for rational points. (However, we shall see a little later in this introduction that the Brauer–Manin obstruction often provides an algorithm to check for rational points.)

How then can one determine whether or not a given Diophantine equation has a solution? One approach is to see if there exist any solutions in larger fields, namely the completions $\mathbb{Q}_p$ and $\mathbb{R}$ of the rationals. This is a simpler problem as one can use the Newton–Raphson method (or it's p-adic analogue, Hensel's lemma) to search for a solution. If one of these fields does not have solutions then there cannot be any rational solutions and the question is answered. Understanding the converse implication, when the existence of solutions in all these larger fields might guarantee a rational solution, is what is known as a local–global problem. The study of such problems in the context of Diophantine equations is the subject of this thesis.

The vanishing of a polynomial defines an algebraic variety and so the above questions can be rephrased in terms of the existence and distribution of rational points on algebraic varieties. This viewpoint allows one to bring to bear the

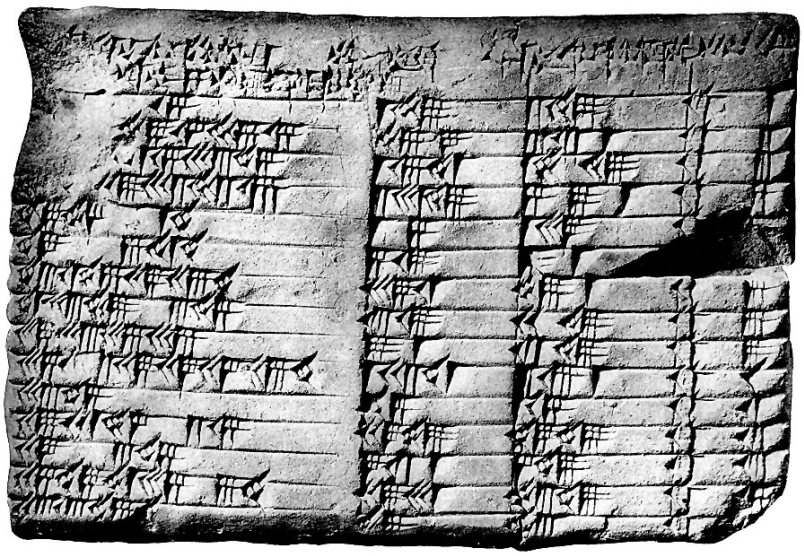

Figure 1.1: Plimpton 322, an ancient Babylonian tablet containing Pythagorean triples. Image source: `https://www.math.ubc.ca/~cass/courses/m446-03/pl322/322.jpg`

tools and techniques of algebraic geometry to shed light on these problems. Given a variety V there is a canonical embedding of the rational points into the product of the local points at each completion

$$\begin{aligned} V(\mathbb{Q}) &\to \prod_{\nu} V(\mathbb{Q}_\nu) \\ x &\mapsto (x, x, \dots). \end{aligned} \tag{1.0.1}$$

One way to consider all these completions at once is via the adeles of $\mathbb{Q}$. This ring is defined as the restricted product of all the completions $\mathbb{Q}_\nu$ with respect to the ring of integers $\mathbb{Z}_\nu$, that is to say it is the subset of the product consisting of elements $(x_\nu)_\nu$ such that each $x_\nu \in \mathbb{Z}_\nu$ for all but finitely many ν. (Note that here we are using the convention that $\mathbb{Z}_\infty = \mathbb{R}$.) The topology on the restricted product is given by a basis of open sets $\prod_\nu U_\nu$ where U_ν is open in $\mathbb{Q}_\nu$ for all ν and $U_\nu = \mathbb{Z}_\nu$ for all but finitely many ν. The reason why the adeles are frequently a more intuitive object to study than the naive product of the local fields is that under this topology the adeles are locally compact and the field $\mathbb{Q}$ forms a discrete cocompact subgroup. Now we may talk about the set of adelic points of a variety $X(\mathbb{A}_\mathbb{Q})$.

Definition 1.0.1. Given a finite set S of places of $\mathbb{Q}$ and $\mathfrak{X}$ an S-integral model for X, the set of adelic points is the restricted product of the sets $X(\mathbb{Q}_\nu)$ with respect to the sets $\mathfrak{X}(\mathbb{Z}_\nu)$.

i.e. An adelic point is a point $(x_\nu)_\nu \in \prod_\nu X(\mathbb{Q}_\nu)$ such that for all but finitely many $\nu \notin S$, x_ν comes from a point in $\mathfrak{X}(\mathbb{Z}_\nu)$.

Remark. This definition is independent of the choice of the set S and the model $\mathfrak{X}$ (see e.g. [93, Section 2.6.3 and Exercise 3.4]).

The adelic points on a variety are a subset of the Cartesian product of the local points. When V is a projective variety, this product coincides with the set of adelic points. We are interested in studying $V(\mathbb{Q})$ and hope to do so by investigating the more tractable set $\prod_\nu V(\mathbb{Q}_\nu)$, but in order for this to be a useful approach we need to understand also the relationship between these two sets.

Definition 1.0.2. Let $\mathcal{V}$ be a class of varieties. We say that the Hasse principle holds for $\mathcal{V}$ if for every $V \in \mathcal{V}$, we have $V(\mathbb{A}_\mathbb{Q}) \neq \varnothing \iff V(\mathbb{Q}) \neq \varnothing$.

This is the simplest kind of density statement for the image of the rational points in the adelic points. We will see later some stronger notions of density (c.f. Definition 1.1.10). The rest of this introduction is devoted to discussing which classes of varieties have this property, and why/how frequently this can fail.

1.1 The Hasse principle and weak approximation

1.1.1 Quadrics

The most well known example of a local-global principle for rational points is the Hasse–Minkowski theorem (e.g. [105, Thm. 8, Ch. IV]).

Theorem 1.1.1. *A quadratic form has non-trivial $\mathbb{Q}$ solution if and only if it has a non-trivial solution over $\mathbb{R}$ and $\mathbb{Q}_p$ for all p.*

Originally proven by Hasse (then extended to number fields by Minkowski) this result is the archetypical local-global statement. In this section we present a (classical) proof of this theorem for quadratic forms in 3 and 4 variables (based heavily on [105]). Already present here are tools and ideas which appear again and again, not just in this thesis, but also in the study of local-global principles in general.

We start with the case of conics. Each conic over $\mathbb{Q}$ is equivalent to an equation of the form $X^2 - aY^2 - bZ^2 = 0$ with a and b squarefree, which we will from here on denote $C(a, b)$. The proof of Hasse–Minkowski for conics proceeds by induction on $m = |a| + |b|$. In the case that $m = 2$, we have $a, b \in \{\pm 1\}$ and it is clear that the conic is everywhere locally soluble exactly when at least one of a and b are $+1$, in which case the conic is also clearly globally soluble. For the induction step, we assume w.l.o.g. $|b| \geqslant |a|$ so that $|b| \geqslant 2$. Note that if $a = 1$ then the conic

$$X^2 - Y^2 = bZ^2$$

always has $\mathbb{Q}$ points regardless of the value of b (for example one may take $(X, Y, Z) = (1, 1, 0)$). Henceforth we assume $a \neq 1$. Let $p \mid b$ and suppose that $p \nmid a$. Then the reduction of $C(a, b)$ to $\mathbb{F}_p$ is defined by the equation

$$X^2 - aY^2 \equiv 0 \bmod p.$$

The conic therefore has a solution in $\mathbb{Q}_p$ if and only if a is a square mod p. Moreover if $a \equiv 0 \bmod p$ then of course it is automatically a square mod p. Hence by the Chinese Remainder Theorem, a is a square mod b and can be written as $a = t^2 - b'b$ for some $|t| \leqslant \frac{|b|}{2}$. The equation $X^2 - aY^2 - b'bZ^2 = 0$ has solution $(t, 1, 1)$ in $\mathbb{Q}$. Suppose that $\widetilde{X}^2 - a\widetilde{Y}^2 - b'\widetilde{Z}^2 = 0$ for some $(\widetilde{X}, \widetilde{Y}, \widetilde{Z})$ in either $\mathbb{Q}$ or $\mathbb{Q}_\nu$. Then

$$\begin{aligned}\left(\widetilde{X}t + a\widetilde{Y}\right)^2 - a\left(\widetilde{Y}t + \widetilde{X}\right)^2 &= \widetilde{X}^2t^2 + a^2\widetilde{Y}^2 - at^2\widetilde{Y}^2 - a\widetilde{X}^2 \\ &= \left(\widetilde{X}^2 - a\widetilde{Y}^2\right)(t^2 - a) \\ &= b\left(b'\widetilde{Z}\right)^2.\end{aligned}$$

Therefore since $t^2 \neq a$ (as a is squarefree and not equal to 1), the existence of a non-trivial solution to $C(a, b)$ is equivalent to the existence of a non-trivial solution to $C(a, b')$ in either $\mathbb{Q}$ or $\mathbb{Q}_\nu$. However $|b'| \leqslant \frac{t^2 - a}{|b|} \leqslant \frac{|b|}{2} + 1$. By the inductive hypothesis $C(a, b')$ is soluble over $\mathbb{Q}$ if and only if it is soluble over all $\mathbb{Q}_\nu$ concluding the proof.

Remark. Essentially what we used at the end was that the conic $C(a, b)$ has a non-trivial solution if and only if b is a norm in the quadratic field extension $\mathbb{Q}(\sqrt{a})/\mathbb{Q}$. Hence the Hasse–Minkowski theorem for conics is equivalent to asking that whenever b is everywhere locally a norm of a quadratic extension, it is also a global norm. This statement is the Hasse norm theorem for quadratic extensions (c.f. Example 1.3.5).

The study of conics is paralleled by the study of quaternion algebras.

Definition 1.1.2. Let K be a field of characteristic 0, $\overline{K}$ the algebraic closure. For $a, b \in K^\times$ the relations

$$i^2 = a, \quad j^2 = b, \quad ij = -ji,$$

define a 4-dimensional k-algebra, with basis $\{1, i, j, ij\}$, which we denote $(a, b)_K$.

When $a = b = -1$ and $K = \mathbb{R}$ this recovers the definition of Hamilton's quaternions $\mathbb{H}$. The connection between the study of quaternion algebras and of quadratic forms is made explicit by the norm form.

Definition 1.1.3. The *norm form* of the algebra $(a, b)_K$ is given by

$$N(x + yi + zj + wij) = x^2 - ay^2 - bz^2 + abw^2.$$

The existence of a nontrivial solution to the conic $C(a,b)$ in K is equivalent to the existence of an element $X+Yi+Zj$ in the quaternion algebra $(a,b)_K$ which is non-zero but has norm 0. The next result tells us when this is possible.

Proposition 1.1.4 ([51, Proposition 1.1.7])**.** *Suppose that K is a local field of characteristic 0 and B a quaternion algebra over K. Then B is either $M_2(K)$, the set of 2×2 matrices in K or B is a division algebra.*

In a division algebra there can be no non-zero elements of norm 0. Hence the conic is soluble over a local field if and only if the associated quaternion algebra is isomorphic to $M_2(K)$, we say the algebra is *split*. If a quaternion algebra is not split, then we say that it is *ramified*. Given a quaternion algebra $(a,b)_{\mathbb{Q}}$ we can obtain an algebra over the local field $\mathbb{Q}_\nu$ by taking the tensor product $(a,b)_{\mathbb{Q}}\otimes_{\mathbb{Q}}\mathbb{Q}_\nu$. Note that if p is an odd prime which does not divide ab, then the reduced conic in $\mathbb{F}_p$ is smooth and thus has a point, so the local quaternion algebra $(a,b)_{\mathbb{Q}}\otimes_{\mathbb{Q}}\mathbb{Q}_\nu$ is split. Hence any quaternion algebra over $\mathbb{Q}$ only ramifies locally at finitely many places. The Hasse-Minkowski theorem for conics can know be rephrased as: the quaternion algebra $(a,b)_{\mathbb{Q}}$ splits if and only if the all of the associated local algebras split. In fact, more than this is true.

Theorem 1.1.5. *There is a one-to-one correspondence between finite, even cardinality sets S of places of $\mathbb{Q}$ and isomorphism classes of quaternion algebras over $\mathbb{Q}$. A quaternion algebra B is associated to the set*

$$S_B=\{\nu : B \text{ is ramified at } \nu\}.$$

Proof. This is a special case of Theorem 1.1.8. □

There is an injective map

$$\{\text{Quaternion algebras}/\mathbb{Q}\}\to\bigoplus_\nu\{\text{Quaternion algebras}/\mathbb{Q}_\nu\},$$

sending a quaternion algebra B to the collection of algebras $(B\otimes_{\mathbb{Q}}\mathbb{Q}_\nu)_\nu$. This is well defined since the local algebras are almost always split. Moreover we can define the map

$$\bigoplus_\nu\{\text{Quaternion algebras}/\mathbb{Q}_\nu\}\xrightarrow{\text{inv}}\mathbb{Z}/2\mathbb{Z},$$

which sends a collection $(B_\nu)_\nu$ to $\sum_\nu\alpha_\nu$ where $\alpha_\nu=0$ if B_ν is split, and $\frac{1}{2}$ otherwise. Again this map is well defined since all but finitely many of the $\alpha_\nu=0$. Theorem 1.1.5 says that quaternion algebras over $\mathbb{Q}$ get mapped to collections which split at an even number of places and hence lie in the kernel of the *inv* map. This means we have a short exact sequence

$$0\to\{\text{Quat. alg's}/\mathbb{Q}\}\to\bigoplus_\nu\{\text{Quat. alg's}/\mathbb{Q}_\nu\}\to\mathbb{Z}/2\mathbb{Z}\to 0.$$

In fact, this sequence is the 2-torsion part of a more general short exact sequence sometimes referred to as the *fundamental short exact sequence of class field theory*. We will see later that this sequence is one of the key building blocks in the Brauer–Manin obstruction to the existence of rational points.

Returning to the Hasse–Minkowski theorem let us look at quadratic forms in 4 variables. We introduce here an invaluable tool in the study of the isotropy of quadratic forms.

Definition 1.1.6. Let k be a local field of characteristic 0 and $a, b \in k^\times$. The Hilbert symbol (over k) of a and b is defined as

$$\left(\frac{a,b}{k}\right) := \begin{cases} 1 \text{ if } X^2 - aY^2 - bZ^2 \text{ has non-trivial } k\text{-solution} \\ -1 \text{ otherwise.} \end{cases}$$

The key property of the Hilbert symbol that makes it a powerful tool for studying the local to global property is the following theorem.

Theorem 1.1.7 (Hilbert's product formula [105, Chapter III, Theorem 3])**.** *For $a, b \in \mathbb{Q}^\times$, we have*

$$\prod_\nu \left(\frac{a,b}{\mathbb{Q}_\nu}\right) = +1.$$

When a and b are distinct primes this recovers the classical quadratic reciprocity formula. One immediate consequence of this theorem is that for $a, b \in \mathbb{Q}^\times$ the number of places ν where $C(a,b)$ has no solution must be even. To apply the Hilbert symbol to our study of quadratic forms in 4 variables we make the observation that every such form can be written as $aX^2 + bY^2 - (cZ^2 + dW^2)$. Assuming that this form is soluble over $\mathbb{Q}_\nu$ means that there must exist $x_\nu \in \mathbb{Q}_\nu$ which is simultaneously represented by both $aX^2 + bY^2$ and $cZ^2 + dW^2$. Equivalently

$$\left(\frac{x_\nu, -ab}{\mathbb{Q}_\nu}\right) = \left(\frac{a,b}{\mathbb{Q}_\nu}\right) \qquad \text{and} \qquad \left(\frac{x_\nu, -cd}{\mathbb{Q}_\nu}\right) = \left(\frac{c,d}{\mathbb{Q}_\nu}\right) \text{ for all } \nu.$$

The heart of the proof of Hasse–Minkowski is that there exist rational numbers whose Hilbert symbols take given values at each place.

Theorem 1.1.8. *Let I a finite set and $(a_i)_{i\in I}$ a collection of elements in $\mathbb{Q}^\times$. Suppose that we are given a collection $(\epsilon_{i,\nu})_{i\in I, \nu\in\Omega_\mathbb{Q}}$ of numbers equal to ± 1. Then there exists a rational number $x \in \mathbb{Q}^\times$ with $\left(\frac{a_i,x}{\mathbb{Q}_\nu}\right) = \epsilon_{i,\nu}$ for all i, ν if and only if the following conditions all hold:*

1. *For almost all ν and i, we have $\epsilon_{i,\nu} = +1$,*
2. *For all $i \in I$, we have $\prod_\nu \epsilon_{i,\nu} = +1$,*
3. *For all ν there exists $x_\nu \in \mathbb{Q}_\nu$ such that $\left(\frac{a_i,x}{\mathbb{Q}_\nu}\right) = \epsilon_{i,\nu}$ for all $i \in I$.*

The necessity of 1 and 3 is immediate, and that of 2 follows from the Hilbert product formula. Sufficiency is a consequence of two very important facts.

Lemma 1.1.9. *Let S be a finite subset of $\Omega_{\mathbb{Q}}$. The image of $\mathbb{Q}$ in $\prod_{\nu\in S}\mathbb{Q}_\nu$ is dense w.r.t the product topology.*

This is a natural extension of the Chinese Remainder Theorem and a special case of a more general local-global property.

Definition 1.1.10. The variety V satisfies *strong approximation* if $V(\mathbb{Q})$ is dense in $V(\mathbb{A}_{\mathbb{Q}})$ with respect to the adelic topology.
V satisfies *weak approximation* if $V(\mathbb{Q})$ is dense in $\prod_\nu V(\mathbb{Q}_\nu)$ with respect to the product topology.

The second big fact we need is that there exist infinitely many primes in any primitive arithmetic progression.

Lemma 1.1.11 (Dirichlet's (weak) theorem on arithmetic progressions)**.** *Let $a\in\mathbb{Z}$ and $q\in\mathbb{N}$ be non-zero, coprime integers. Then there are infinitely many primes p such that $p\equiv a \bmod q$.*

Of course the proof of this is suppressed; the reader is referred to [41] for more details.

Proof of Theorem 1.1.8. We may assume that all the a_i are integers. Let S be the set of places consisting of $2,\infty$ and all prime factors of the a_i. Let T be the set of places such that there exists an $i\in I$ with $\epsilon_{i,\nu}=-1$. There are now two cases to consider:
Case 1: $S\cap T=\varnothing$ Define

$$a=\prod_{\substack{\ell\in T\\ \ell\neq\infty}}\ell \qquad \text{and} \qquad q=8\prod_{\substack{\ell\in S\\ \ell\neq 2,\infty}}\ell.$$

Since the sets S and T are disjoint, the integers a and q must be coprime. Dirichlet's theorem then gives us a prime p which is not in $S\cup T$ and which is congruent to a modulo q. Set $x=ap$. For all $\nu\in S$ we know that $\epsilon_{i,\nu}=+1$. When $\nu=\infty$, since x is positive it follows that $\left(\frac{a_i,x}{\mathbb{Q}_\nu}\right)=+1=\epsilon_{i,\nu}$. For every finite place ν in S, we know that $x\equiv a^2 \bmod m$ and hence x is a square in $\mathbb{Q}_\nu$. At finite places ℓ not in $T\cup\{p\}$, x is an ℓ-adic unit and thus $\left(\frac{a_i,x}{\mathbb{Q}_\ell}\right)=+1=\epsilon_{i,\ell}$. For $\ell\in T$ since $\ell\mid x$ the conic $C(a_i,x)$ reduces to $a_iX^2=Z^2 \bmod \ell$. This has a non-trivial solution if and only if a_i is a square in $\mathbb{Q}_\ell$ but since $\epsilon_{i,\ell}=-1$, this would violate condition 3. Therefore $\left(\frac{a_i,x}{\mathbb{Q}_\ell}\right)=-1$. Finally consider $\nu=p$, the product formula gives us

$$\left(\frac{a_i,x}{\mathbb{Q}_p}\right)=\prod_{\nu\neq p}\left(\frac{a_i,x}{\mathbb{Q}_\nu}\right)=\prod_{\nu\neq p}\epsilon_{i,\nu}=\epsilon_{i,p}.$$

Case 1: General S and T We first observe that the squares in $\mathbb{Q}_\nu^\times$ form an open subset (this is pretty obvious for $\mathbb{R}$ and for $\mathbb{Q}_p$ it follows from squares being determined by their residue class). Using Lemma 1.1.9, we can find an $x'\in\mathbb{Q}^\times$

such that x'/x_ν is square in $\mathbb{Q}_\nu^\times$ for all $\nu \in S$. Therefore $\left(\frac{a_i,x'}{\mathbb{Q}_\nu}\right) = \left(\frac{a_i,x_\nu}{\mathbb{Q}_\nu}\right) = \epsilon_{i,\nu}$. We can now apply Case 1 above with the new constants $\epsilon_{i,\nu}(a_i,x')_\nu$ which are $+1$ at all places in S. This yields $y \in \mathbb{Q}^\times$ with $\left(\frac{a_i,y}{\mathbb{Q}_\nu}\right) = \epsilon_{i,\nu}\left(\frac{a_i,x'}{\mathbb{Q}_\nu}\right)$ for all i and ν. Setting $x = x'y$ completes the proof. □

By the Hilbert product formula, we know that $\prod_\nu \left(\frac{a,b}{\mathbb{Q}_\nu}\right) = \prod_\nu \left(\frac{c,d}{\mathbb{Q}_\nu}\right) = +1$ and moreover we know that there exists $x_\nu \in \mathbb{Q}_\nu^\times$ such that $\left(\frac{x_\nu,-ab}{\mathbb{Q}_\nu}\right) = \left(\frac{a,b}{\mathbb{Q}_\nu}\right)$ and $\left(\frac{x_\nu,-cd}{\mathbb{Q}_\nu}\right) = \left(\frac{c,d}{\mathbb{Q}_\nu}\right)$. Therefore we may apply Theorem 1.1.8 with

$$a_1 = -ab, a_2 = -cd, \epsilon_{1,\nu} = \left(\frac{a,b}{\mathbb{Q}_\nu}\right) \text{ and } \epsilon_{2,\nu} = \left(\frac{c,d}{\mathbb{Q}_\nu}\right),$$

to conclude that there exists $x \in \mathbb{Q}^\times$ with

$$\left(\frac{x,-b}{\mathbb{Q}_\nu}\right) = \left(\frac{a,b}{\mathbb{Q}_\nu}\right) \text{ and } \left(\frac{x,-cd}{\mathbb{Q}_\nu}\right) = \left(\frac{c,d}{\mathbb{Q}_\nu}\right),$$

for every ν. This means that the conics $aX^2 + bY^2 = xZ^2$ and $cX'^2 + dY'^2 = xZ'^2$ are both soluble in every $\mathbb{Q}_\nu$. We can thus conclude from the previous discussion that both of these conics are $\mathbb{Q}$-soluble implying Hasse–Minkowski for $n = 4$.

Remarks. (i) We have seen above that the solubility of conics at least is intimately connected to the arithmetic of certain algebras. This is a theme which will be expanded upon in the sequel.

(ii) In the second proof, our approach was to study the quadric surface $f := aX^2+bY^2+cZ^2+dW^2 = 0$ by looking at the solutions of the pair of forms $(f_1, f_2)_x := (aX^2 + bY^2 - xT^2, cZ^2 + dW^2 - x\widetilde{T}^2)$. We are parametrising the solutions to f into a 1-parameter family as x varies. The base of the fibration is $\mathbb{A}^1$ where we know that the rational points are dense in the adelic points and each fibre is a conic where we know that the local-global principle holds. This is the heart of *the fibration method*, which we shall explore further in Chapter 2.

(iii) In the fibration argument above, we need to appeal to the infinitude of primes in arithmetic progressions. Here we see the first instance in this thesis of the distribution of the zeroes of L-functions playing a role in Diophantine geometry. This is just the tip of the iceberg.

(iv) By the above theorem all quadrics satisfy the Hasse principle. In fact not only that but quadrics with points everywhere locally also satisfy weak approximation. We reserve the proof till Chapter 2.

One particularly nice consequence of the Hasse–Minkowski theorem is that investigating the existence of rational points on quadrics becomes a finite computation. Indeed the Lang–Weil bounds guarantee that for large primes p, the

quadric always has a smooth $\mathbb{F}_p$-point (which may be lifted by Hensel's lemma to a $\mathbb{Q}_p$-point) so it remains to check the real place and some finite number of primes. This gives us hope to get around the difficulty of undecidability discussed before.

1.1.2 Hasse principle failures

Alas in degrees higher than 2 the picture is less rosy. The first counter-example to the Hasse principle was provided by Lind [75] and Reichardt [96] independently.

Example 1.1.12. The curve defined by $x^4 - 17y^4 = 2z^2$ does not satisfy the Hasse principle.

Proof. It is clear that the curve has $\mathbb{R}$ points. Moreover if p is a prime such that the reduced curve is smooth over $\mathbb{F}_p$ (i.e. $p \neq 2$ or 17) then we can apply the Hasse-Weil bound to deduce the existence of smooth $\mathbb{F}_p$-point. Hence via Hensel's lemma there is a $\mathbb{Q}_p$-point. Finally note that the reduced equation $x^4 - 17y^4 \equiv 0 \mod 64$ has solution $(x, y) = (3, 1)$ which can be lifted to a $\mathbb{Q}_2$ solution to the original solution. Likewise, $2z^2 - x^4 \equiv 0 \bmod 17$ has solution $(x, z) = (6, 1)$. Therefore solutions exist over every $\mathbb{Q}_\nu$.

Now suppose that there were a rational solution (x, y, z). By multiplying denominators through, we may assume w.l.o.g. that x, y, z are integers and x and z are coprime. Suppose that p is an odd prime dividing y, then $x^4 \equiv 17y^4 \bmod p$ and so 17 must be a square mod p. Thus by quadratic reciprocity, p is a square mod 17, as is 2 and -1 incidentally. Therefore y is a product of squares mod 17, making it a square itself. Writing $y = w^2$ we have that $x^4 \equiv 2w^4 \bmod 17$. From this we conclude that 2 is fourth power modulo 17 which is not true and hence no rational solution exists. □

The following famous example due to Selmer shows that the Hasse principle even fails in degree 3.

Example 1.1.13 (Selmer 1951 [104])**.** The cubic curve defined by the equation $3x^3 + 4y^3 + 5z^3 = 0$ does not satisfy the Hasse principle.

Of course the existence of Hasse principle failures means that weak approximation can also fail but the following example demonstrates that weak approximation can fail even when the variety has a rational point.

Example 1.1.14 (Colliot-Thélène–Sansuc 1979[34])**.** The variety defined by the smooth intersection of two quadrics in $\mathbb{P}^5$ given by the equations

$$\begin{cases} x^2 + y^2 + z^2 = uv \\ x^2 + 2y^2 + t^2 = (u - v)(u - 2v) \end{cases}$$

has rational points but fails weak approximation.

The existence of classes of varieties which need not satisfy the Hasse principle leads one to wonder a few things.

1. Why does the Hasse principle fail?

2. If the Hasse principle can fail for a class of varieties, how common is this failure?

The ideas used to attack the former question are the subject of Section 1.2 and the latter Section 1.3.

1.2 The Brauer–Manin obstruction

Manin addressed Question 1 in his talk at the 1970 ICM in Nice [84]. He proposed an explanation for failures of the Hasse principle by constructing an obstruction set, denoted $V(\mathbb{A}_{\mathbb{Q}})^{\mathrm{Br}}$, so that

$$V(\mathbb{Q}) \subset V(\mathbb{A}_{\mathbb{Q}})^{\mathrm{Br}} \subset V(\mathbb{A}_{\mathbb{Q}}).$$

When $V(\mathbb{A}_{\mathbb{Q}}) \neq \varnothing$ but $V(\mathbb{A}_{\mathbb{Q}})^{\mathrm{Br}} = \varnothing$ then it follows that $V(\mathbb{Q}) = \varnothing$ and we say that the Hasse principle is *obstructed* by the set $V(\mathbb{A}_{\mathbb{Q}})^{\mathrm{Br}}$. The explicit construction of this set makes use of the Brauer group associated to the variety V and so we take a brief detour to introduce it.

1.2.1 The Brauer group of a field

The development of class field theory in the late 19th and early 20th century was motivated by the problem of classifying the abelian extensions of a number field. A natural follow up is to ask what happens if one considers not field extensions but algebras defined over k which are allowed to be noncommutative. One property which it seem natural to ask for in the generalisation of field extensions to algebras is the following.

Definition 1.2.1. A k-algebra A is *simple* if the only (two-sided) ideals of A are (0) and A itself.

Of course a field extension satisfies this property but so too does, for example, the set of Hamiltonians $\mathbb{H}$ over $\mathbb{R}$, which are non-commutative. To classify finite dimensional simple algebras we can use Wedderburn's theorem.

Theorem 1.2.2 ([51, Theorem 2.1.3])**.** *Let A be a finite dimensional simple algebra over a field k. Then there exist an integer $n \geqslant 1$ and a division algebra $D \supset k$ so that A is isomorphic to the matrix ring $M_n(D)$. Moreover, the division algebra D is uniquely determined up to isomorphism.*

A *central* algebra is a k-algebra A for which the center (i.e. the set of all elements $x \in A$ such that $x \cdot y = y \cdot x$ for each $y \in A$) is just k.

Example 1.2.3. Quaternion algebras are central simple algebras. Indeed let A a quaternion algebra over k and suppose that $x = \alpha + \beta i + \gamma j + \delta k$ is in the center of A. Then

$$0 = jx - xj = 2k(\beta + \delta j).$$

Since k is invertible in A, we must have $\beta = \delta = 0$. Similarly

$$0 = (\alpha + \beta i)j - j(\alpha + \beta i) = 2k\beta,$$

and thus $\beta = 0$ and so $x \in k$. This shows the centrality of A. To see that it is simple suppose that I is a double-sided, non-empty ideal of A. Let $y = a + bi + cj + dk$ a non-zero element of I, we may assume w.l.o.g. $a \neq 0$. Since $yj - jy \in I$ and $2k$ is in invertible in A it must be the case that $b + dj$ and $bi + dk$ are in I, similarly $a+bi$ and $a+dk$ are in I. Therefore $y-(a+bi)-(a+cj)-(a+dk) = -2a$ is an element of I hence I contains all elements of A. Thus every non-zero double sided ideal of A is A itself.

Including the notion of centrality allows the following refined classification.

Theorem 1.2.4 ([51, Theorem 2.2.1])**.** *Let k be a field and A a finite dimensional k-algebra. Then A is a central simple algebra if and only if there exist an integer $n > 0$ and a finite field extension K/k so that $A \otimes_k K$ is isomorphic to the matrix ring $M_n(K)$.*

Central simple algebras, henceforth abbreviated to CSA, then are precisely the ones which admit a notion of being *split*, analogous to the splitting of quaternion algebras which carried so much arithmetic content. Given a finite extension K/k, we could ask: which central simple k-algebras are split by K?

Lemma 1.2.5. *If A and B are central simple k-algebras then so is $A \otimes_k B$.*

Proof. The lemma is a simple consequence of Theorem 1.2.4. Choose a large enough extension K/k so that both A and B are split by K. Then the natural isomorphism

$$(A \otimes_k K) \otimes_K (B \otimes_k K) \cong (A \otimes_k B) \otimes_k K$$

reduces the lemma to showing that

$$M_n(K) \otimes_K M_m(K) \cong M_{mn}(K),$$

which is straightforward. □

In particular we could map the central simple algebras of dim n into those of dim mn by simply tensoring by the matrix algebra $M_m(k)$. In fact, this defines an embedding.

Lemma 1.2.6. *The map $A \mapsto A \otimes_k M_m(k)$ is an injection from CSAs of dimension n to CSA's of dimension mn.*

Proof. Let A and B be CSA's such that $A \otimes_k M_m(k) \cong B \otimes_k M_m(k)$. By the existence part of Theorem 1.2.2, there exists a division algebra D such that $A \cong M_a(D)$ for some $a \in \mathbb{N}$ and therefore $A \otimes_k M_n(k)$ is a matrix algebra over D. By the uniqueness part of Theorem 1.2.2, $B \otimes M_n(D)$ must also be a matrix algebra over D. Therefore $B \cong M_b(D)$ for some $b \in \mathbb{N}$ and since A and B have the same dimension we conclude that they are isomorphic. □

This motivates the definition of an equivalence relation on central simple algebras. Namely two algebras A and B are equivalent if there exist $m, n \geqslant 1$ such that $A \otimes M_m(k) \cong B \otimes M_n(k)$.

Corollary 1.2.7. *Two equivalent CSAs of the same dimension are isomorphic.*

This means that together the dimension and the equivalence class of these algebras contains all the information we could possibly want about them.

Definition 1.2.8. The set of equivalence classes of CSAs over k, equipped with the operation of tensor product, forms an abelian group which we denote $\mathrm{Br}(k)$.

Example 1.2.9.

1. $\mathbb{C}$: Let D be a CSA over $\mathbb{C}$. Then the norm map is a polynomial over $\mathbb{C}$ with non-trivial solution. Therefore D is not a division algebra and hence is split. Therefore $D \cong M_n(\mathbb{C})$ for some n.

2. $\mathbb{R}$: By Theorem 1.2.2, we may write every CSA as $M_n(D)$ for some division algebra $\mathbb{R} \subset D$. By a well-known theorem of Frobenius the finite-dimensional associative division algebras over $\mathbb{R}$ are isomorphic to either $\mathbb{R}$, $\mathbb{H}$ or $\mathbb{C}$. Hence every CSA is either equivalent to $\mathbb{R}$ or $\mathbb{H}$.

Proposition 1.2.10 ([51, Corollary 4.4.8])**.** *Br(k) is an abelian group. Every element has finite order.*

The Hasse–Minkowski theorem for conics gave us a short exact sequence connecting the quaternion algebras over $\mathbb{Q}$ with collections of quaternion algebras over all completions. One of the most pivotal facts in class field theory and in the understanding of local-global principles in general is that this fact generalises.

Theorem 1.2.11.

1. *For every place ν of $\mathbb{Q}$, there is an injective map* $\mathrm{inv}_\nu : Br(\mathbb{Q}_\nu) \to \mathbb{Q}/\mathbb{Z}$ *with image*

$$\begin{cases} \mathbb{Z}/2\mathbb{Z} \text{ if } \nu = \infty \\ \mathbb{Q}/\mathbb{Z} \text{ if } \nu = p. \end{cases}$$

2. *The following sequence is exact*

$$0 \to Br(\mathbb{Q}) \to \bigoplus_\nu Br(\mathbb{Q}_\nu) \to \mathbb{Q}/\mathbb{Z} \to 0.$$

This is a beautiful local-global statement for the splitting of certain algebras over a field. But we are interested in local-global principles as they pertain to solving equations. What can we build on from this theory? In the next section we will study the Brauer group associated to the function field of a variety and see what it can tell us about the local-global property for the existence of rational points.

1.2.2 The Brauer group of a variety

It is a common refrain in algebraic geometry that to understand the geometry of a scheme one should look at its rational functions. Given a variety V/k, we could therefore ask about the Brauer group, as defined in the previous section, of the function field $k(V)$. However it is important that we only focus on those functions which are well-defined everywhere on V.

Example 1.2.12 (inspired by [14])**.** If $V = \mathbb{A}^1_{\mathbb{Q}}$ then the function field is $\mathbb{Q}(V) = \mathbb{Q}(t)$. The Brauer group is quite large but two examples of quaternion algebras which live in this group are given by

$$\mathcal{A} = (2,3)_{\mathbb{Q}(t)} \qquad \text{and} \qquad \mathcal{B} = (-1,t)_{\mathbb{Q}(t)}.$$

The first of these is in fact an element of $\mathrm{Br}(\mathbb{Q})$ and the functions defining this algebra are constant, so they are indeed defined everywhere on X. However the second algebra doesn't make sense at the point $t = 0$ where the algebra $\mathcal{B}$ is not well-defined!

Instead of considering the full field of rational functions then, we focus on those algebras which are made up of rational functions that are well defined above a point. This means defining the Brauer group of the local ring $\mathcal{O}_{V,P}$ for a point $P \in V$.

Definition 1.2.13. Let R be a local ring with maximal ideal $\mathfrak{m}$. An *Azumaya algebra* over R is a free R-algebra A such that $A/\mathfrak{m}A$ is a CSA over the field $R/\mathfrak{m}$.

Analogously to the case of CSAs, we say two Azumaya algebras A and B over R are equivalent if there exists $m, n > 0$ such that $A \otimes_R M_m(R) \cong B \otimes_R M_n(R)$.

Definition 1.2.14. The Brauer group of R, $\mathrm{Br}(R)$, is the set of all equivalence classes of Azumaya algebras over R under the operation of tensor product.

This now allows us to form a notion of the Brauer group for a variety.

Definition 1.2.15. Let V be a smooth variety over k. The Brauer group of V is defined as

$$\mathrm{Br}(V) := \bigcap_P \mathrm{Br}(\mathcal{O}_{V,P}),$$

where the intersection runs over all closed points of V.

Remark. Since $\mathcal{O}_k \subset \mathcal{O}_{V,P}$ for every $P \in V(k)$ we have that any CSA over k lives in $\mathrm{Br}(\mathcal{O}_{V,P})$ for each $P \in V(k)$ and therefore in $\mathrm{Br}(V)$.

We should note that from this definition it is unclear that this group is explicitly computable in practical terms. There is an alternative definition to which one can bring to bear the full power of cohomological methods.

Definition 1.2.16 (Cohomological Version)**.** Let V be a quasi-compact scheme. Then the cohomological Brauer group is defined as

$$\widetilde{\mathrm{Br}}(V) := \mathrm{H}^2_{\acute{e}t}(V, \mathbb{G}_m)_{\mathrm{tors}}.$$

When V is a quasi-compact, quasi-projective, seperated scheme, this coincides with the previous definition (this is a theorem of Gabber [50]). Since $\mathrm{Br}(V) \subset \mathrm{Br}(k(V))$, the algebras in the Brauer group act like regular functions on V and hence we can evaluate them at a point to get a class in $\mathrm{Br}(k)$. For a given Brauer group element $\mathcal{A} \in \mathrm{Br}(V)$ we will denote this map from $V(k)$ to $\mathrm{Br}(k)$ by $\mathrm{ev}_{\mathcal{A}}$. This along with the fundamental short exact sequence of class field theory yields the following commutative diagram:

$$\begin{array}{ccccccccc} & & V(\mathbb{Q}) & \longrightarrow & V(\mathbb{A}_{\mathbb{Q}}) & & & & \\ & & \downarrow{\scriptstyle \mathrm{ev}_{\mathcal{A}}} & & \downarrow{\scriptstyle \mathrm{ev}_{\mathcal{A}}} & & & & \\ 0 & \longrightarrow & \mathrm{Br}(\mathbb{Q}) & \longrightarrow & \bigoplus_{\nu} \mathrm{Br}(\mathbb{Q}_{\nu}) & \longrightarrow & \mathbb{Q}/\mathbb{Z} & \longrightarrow & 0. \end{array}$$

For each $\mathcal{A} \in \mathrm{Br}(V)$ the above diagram gives a map $V(\mathbb{A}_{\mathbb{Q}}) \to \mathbb{Q}/\mathbb{Z}$. Therefore we have a pairing

$$\langle \cdot, \cdot \rangle : V(\mathbb{A}_{\mathbb{Q}}) \times \mathrm{Br}(V) \to \mathbb{Q}/\mathbb{Z}.$$

The Brauer–Manin set is the collection of adelic points orthogonal to the Brauer group under this pairing

$$V(\mathbb{A}_{\mathbb{Q}})^{\mathrm{Br}} := \{(P_{\nu})_{\nu} \in V(\mathbb{A}_{\mathbb{Q}}) :< P, \mathcal{A} >= 0 \, \forall \mathcal{A} \in \mathrm{Br}(V)\}.$$

That this is a subset of the adelic points is a tautology and the commutativity of the diagram demonstrates that the image of $V(\mathbb{Q})$ in $V(\mathbb{A}_{\mathbb{Q}})$ is a subset of the Brauer–Manin set, hence this aligns with the idea of an obstruction set introduced at the start of this section.

Definition 1.2.17. We say that the Brauer–Manin obstruction is the only obstruction to the Hasse principle if $V(\mathbb{A}_{\mathbb{Q}})^{\mathrm{Br}} \neq \varnothing \implies V(\mathbb{Q}) \neq \varnothing$. Similarly, if $\overline{V(\mathbb{Q})} = V(\mathbb{A}_{\mathbb{Q}})^{\mathrm{Br}}$ then we say that the Brauer–Manin obstruction is the only obstruction to weak approximation.

Definition 1.2.18. We say that there is no Brauer–Manin obstruction to the Hasse principle, or that the obstruction is empty, if $V(\mathbb{A}_{\mathbb{Q}})^{\mathrm{Br}} \neq \varnothing$. Similarly, if $V(\mathbb{A}_{\mathbb{Q}})^{\mathrm{Br}} = V(\mathbb{A}_{\mathbb{Q}})$ we say that there is no Brauer–Manin obstruction to weak approximation.

Example 1.2.19. Consider the projective variety which is a smooth projective model of the affine variety defined by the equation

$$X^2 + Y^2 = (3 - T^2)(T^2 - 2).$$

By [35, Thm. B], the quotient $\mathrm{Br}(V)/\mathrm{Br}(\mathbb{Q}) \cong \mathbb{Z}/2\mathbb{Z}$ is generated by the quaternion algebra $(-1, T^2 - 2)_{\mathbb{Q}}$. This means that the Brauer–Manin set is given by

$$\left\{(x_\nu, y_\nu; t_\nu)_\nu \in V(\mathbb{A}_{\mathbb{Q}}) : \prod_\nu \left(\frac{-1, t_\nu^2 - 2}{\mathbb{Q}_\nu}\right) = +1\right\}.$$

There are four different types of places: the real place, finite primes $p \equiv \pm 1 \bmod 4$ and the prime 2.

- Since $(3-t_\infty^2)(t_\infty^2-2)$ is a sum of squares in $\mathbb{R}$ both brackets must have the same sign. They cannot both be negative (because $t_\infty^2 > 3 \implies t_\infty^2 - 2 > 0$) therefore both brackets are positive and hence $\left(\frac{-1, t_\infty^2 - 2}{\mathbb{R}}\right) = +1$.
- At odd primes $p \equiv 1 \bmod 4$ every element is a sum of two squares and therefore $\left(\frac{-1, t_p^2 - 2}{\mathbb{Q}_p}\right) = +1$.
- At primes $p \equiv 3 \bmod 4$ we have $\left(\frac{-1, t_p^2 - 2}{\mathbb{Q}_p}\right) = (-1)^{v_p(t_p^2 - 2)}$. Since we know that $(3 - t_p^2)(t_p^2 - 2)$ is a sum of two squares in $\mathbb{Q}_p$, we must have that $v_p(3 - t_p^2) \equiv v_p(t_p^2 - 2) \bmod 2$. But there is no way that a prime could divide both $3 - t_p^2$ and $1 - (3 - t_p^2) = t_p^2 - 2$, therefore we must have $\left(\frac{-1, t_p^2 - 2}{\mathbb{Q}_p}\right) = +1$.
- Lastly at the prime 2, we have $\left(\frac{-1, t_2^2 - 2}{\mathbb{Q}_2}\right) = (-1)^{\frac{(t_2^2 - 2) - 1}{2}} = -1$.

Therefore we conclude that for any collection of local solutions $(x_\nu, y_\nu; t_\nu)_\nu$ we have $\prod_\nu \left(\frac{-1, t_\nu^2 - 2}{\mathbb{Q}_\nu}\right) = -1$ and thus the Brauer–Manin set is empty, obstructing the existence of rational points.

Remark. Similar computations to those above are performed in Chapter 4 where we study how frequently the Hasse principle fails in a family of such surfaces.

This example was first shown to fail the Hasse principle by Iskovskikh [66] who, ironically, suggested it as an example of a variety whose lack of rational points was not explained by Manin's construction. Instead the community would have to wait nearly three decades before Skorobogatov gave the first such example in [111]. Understanding for which classes of varieties the Brauer–Manin obstruction is the only one is one of the most important topics in Diophantine geometry. We are guided by the following well-trusted conjecture.

Conjecture 1.2.20 ([29])**.** *Suppose that V is a smooth, proper, rationally connected and geometrically irreducible variety. Then the Brauer–Manin obstruction is the only one.*

We will see in Chapter 2 how some of the ideas involved in the proof of the Hasse–Minkowski theorem can be expanded upon to prove this conjecture for certain classes of varieties. Often fully determining the Brauer group is still an arduous task. In Chapter 7, we find exactly one element in the Brauer group of the variety described below as part of our investigation into weak approximation in that setting.

Theorem A. *Let $V \subset \mathbb{P}^2 \times \mathbb{P}^3$ be a smooth projective model for the variety defined by the equation*

$$xyt_1^2 + xzt_2^2 + yzt_3^2 + F(x,y,z)t_4^2 = 0,$$

where F is a positive definite ternary quadratic form. If F satisfies $F(0,y,z) \in \mathbb{Q}^2$ for all $y,z \in \mathbb{Q}$ then weak approximation holds for V.

1.3 Families of varieties

One need not restrict attention to a single variety. It is equally interesting to look at a family of varieties and ask which varieties in that family have points or satisfy a local-global principle. All of them, none or only some? In the latter case, how many of the varieties have this property? And how are they distributed?

Specifically consider a projective variety Y defined over $\mathbb{Q}$ equipped with a dominant morphism $\pi : Y \to X$, where X is a smooth projective variety over $\mathbb{Q}$ and the generic fibre of π is geometrically integral. For some choice of height function H, we define the counting functions

$$\begin{aligned} N(X,B) &= \#\{x \in X(\mathbb{Q}) : H(x) \leqslant B\}, \\ N_{\mathrm{loc}}(X,B,\pi) &= \#\{x \in X(\mathbb{Q}) : H(x) \leqslant B \text{ and } \pi^{-1}(x)\,(\mathbb{A}_{\mathbb{Q}}) \neq \varnothing\}, \\ N_{\mathrm{glob}}(X,B,\pi) &= \#\{x \in X(\mathbb{Q}) : H(x) \leqslant B \text{ and } \pi^{-1}(x)\,(\mathbb{Q}) \neq \varnothing\}. \end{aligned}$$

Understanding the ratio $N_{\mathrm{loc}}(X,B,\pi)/N(X,B)$ as $B \to \infty$ tells us the density of varieties in the family defined by π which have local solutions everywhere. Similarly, $N_{\mathrm{glob}}(X,B,\pi)/N(X,B)$ determines the proportion of varieties in the family which have rational solutions and counter-examples to the Hasse principle are counted by $N_{\mathrm{glob}}(X,B,\pi)/N_{\mathrm{loc}}(X,B,\pi)$. These ratios have been studied for a wide array of families in the literature. One of the most famous is the paper of Poonen and Voloch [94]. They consider the case $X = \mathbb{P}^m$ where $m = \binom{n+d}{d}$ and the fibration defined by

$$Y : \left\{ \sum_{(x_0,\dots,x_n) \in \mathbb{Z}^{n+1}_{\mathrm{prim}}} \sum_{i_0+\cdots+i_n=d} \sum_{(a_{0,i_0},\dots,a_{n,i_n}) \in \mathbb{Z}^m_{\mathrm{prim}}} a_{0,i_0} \dots a_{n,i_n} x_0^{i_0} \dots x_n^{i_n} = 0 \right\}$$

$$\begin{aligned} \pi : Y &\to \mathbb{P}^m \\ (x_0,\dots,x_n; a_{0,i_0} \dots a_{n,i_n}) &\mapsto (a_{0,i_0} \dots a_{n,i_n}). \end{aligned}$$

The variety Y here parametrises all homogeneous degree d polynomials in $n+1$ variables and the map π projects onto the coefficients of the forms. Hence each fibre is some degree d form in $n+1$ variables and the frequency with which fibers are soluble corresponds exactly to the frequency with which such polynomials admit solutions. Poonen and Voloch showed that if $n, d \geqslant 2$ and $(n,d) \neq (2,2)$ then in the limit as $B \to \infty$ we have

$$N_{\mathrm{loc}}(\mathbb{P}^m, B, \pi)/N(\mathbb{P}^m, B) \to c \tag{1.3.1}$$

for some constant c which is expressed as a product of factors c_ν at each place ν of $\mathbb{Q}$. Moreover under the assumption of Conjecture 1.2.20 they showed that for $2 \leqslant d \leqslant n$ we have

$$\lim_{B\to\infty} \left(N_{\mathrm{loc}}(\mathbb{P}^m, B, \pi) - N_{\mathrm{glob}}(\mathbb{P}^m, B, \pi)\right)/N(\mathbb{P}^m, B) = 0.$$

Based on this and some heuristic computations they made the following conjectures.

Conjecture 1.3.1. *As $B \to \infty$ we have*

1. *If $d > n+1$ then $N_{glob}(\mathbb{P}^m, B, \pi)/N(\mathbb{P}^m, B) \to 0$.*
2. *If $d < n+1$ then $N_{glob}(\mathbb{P}^m, B, \pi)/N(\mathbb{P}^m, B) \to c$, where $c = \prod_\nu c_\nu$ as above.*

A recent preprint of Le Boudec–Browning–Sawin [11] establishes part of this conjecture, namely they show that 100% of hypersurfaces with $n \geqslant d$ satisfy the Hasse principle, except possibly in the case $n = d = 3$.

Notably there are two regimes where these conjectures do not give a prediction, the cases when $d = n+1$ and when $(d,n) = (2,2)$. A special case of the former regime is given by the result of Bhargava discussed below.

Example 1.3.2. Bhargava [8] showed that the proportion of planar cubic curves (ordered by the height of the coefficients) which have everywhere local points but no $\mathbb{Q}$-points is at least 28%. Moreover he conjectured that the proportion of planar cubic curves with everywhere local points which have no global points is $\frac{2}{3}$. These results are complimented by the work of Bhargava–Cremona–Fisher [9] who showed that the proportion of planar cubic curves with a point locally everywhere is $\approx 97.256\%$.

This gives us an essentially complete picture of the case $(d,n) = (3,2)$. However the next simplest case in this regime $(d,n) = (4,3)$ is still completely open. We do not even have a prediction for what to expect in this case. The other case not covered by the conjecture are the plane conics where $(d,n) = (2,2)$. Here Serre has provided an upper bound [108] and Hooley a lower bound [64] for the number which are everywhere locally soluble, both of which are of the order of magnitude $B^6/(\log B)^{\frac{1}{2}}$. In Chapter 6, we study the problem of everywhere locally soluble plane conics in the special case that they are diagonal (more details below).

1.3.1 The conjecture of Loughran and Smeets

The other major conjecture in this setting is that of Loughran–Smeets. In their paper [78], Loughran and Smeets studied families defined by proper smooth irreducible varieties with a dominant map to projective space that has geometrically integral generic fibre. To study the frequency with which the fibres in such a family are everywhere locally soluble they introduced an invariant of the family which we describe now briefly.

Suppose $D \in \mathbb{P}^{n(1)}$ is a codimension 1 point then the absolute Galois group $Gal(\overline{\kappa(D)}, \kappa(D))$ acts on the irreducible components of $\pi^{-1}(D) \otimes \overline{\kappa(D)}$. Pick $\Gamma_D(\pi)$ a finite subgroup through which the action factors and define $\delta_D(\pi) = \#\Gamma_D(\pi)^o / \#\Gamma_D(\pi)$ where $\Gamma_D(\pi)^o$ is the set of those $\gamma \in \Gamma_D(\pi)$ which fixes some multiplicity 1 geometrically irreducible component of $\pi^{-1}(D)$. Then we define $\Delta(\pi) = \sum_{D \in \mathbb{P}^{n(1)}} (1 - \delta_D(\pi))$.

Definition 1.3.3. A scheme X of finite type over a perfect field k is called *split* if it contains an irreducible component of multiplicity 1 which is geometrically irreducible. Moreover it is *pseudo-split* if every element of $\mathrm{Gal}(\overline{k}/k)$ fixes some irreducible component of $X \times_k \overline{k}$ of multiplicity 1.

Note that if $\pi^{-1}(D)$ is split then $\delta_D(\pi) = 1$ since by definition there is an irreducible component which is fixed by all elements of the Galois group. Hence in the sum defining $\Delta(\pi)$ we need only include the codimension one points with non-split fibres. Since we assume that the generic fibre is geometrically integral (and thus split) it follows that $\Delta(\pi)$ is in fact a finite sum.

Example 1.3.4. Consider the Poonen–Voloch family defined above and suppose $n, d \geqslant 2$ and $(n, d) \neq (2, 2)$. In this case, every fibre above a codimension one point is geometrically irreducible, hence split, and so $\Delta_D(\pi) = 0$.

The main result of [78] showed that

$$N_{\mathrm{loc}}(\mathbb{P}^n, B, \pi)/N(\mathbb{P}^n, B) \ll (\log B)^{-\Delta(\pi)},$$

meaning as soon as this invariant $\Delta(\pi)$ is non-zero we should expect most fibers in the family to fail to be everywhere locally soluble. The exact same construction allows one to define the invariant $\Delta(\pi)$ for any family fibred over a non-singular proper projective variety. In [19, Theorem 1.10], Loughran and Smeets' upper bound was replicated in the case that the base is a quadric hypersurface of dimension at least 3. In Chapter 5, we extend this result to families over any smooth hypersurface X of sufficiently high dimension.

Theorem B. *Let $F \in \mathbb{Z}[x_0, \ldots, x_n]$ be a non-singular homogeneous degree d form and X the smooth projective hypersurface which it defines. Let H be the height function defined by $H(x) = \max_i |x_i|$ for $(x_0, \ldots, x_n) \in \mathbb{Z}^{n+1}_{prim}$ such that $[x] = (x_0 : \cdots : x_n)$. Suppose that $\pi : Y \to X$ is a dominant map with geometrically integral generic fibre from Y a smooth projective variety. If $n > 2^d(d-1)$ then we have*

$$N_{loc}(X, B, \pi) \ll \frac{B^{n+1-d}}{(\log B)^{\Delta(\pi)}}.$$

In [78], the authors conjectured that under certain conditions their upper bounds should be sharp. The first condition is that the set of adelic points is non-empty. The necessity of this is clear. The second condition is that there is no fibre above which every irreducible component has multiplicity 2 or higher. This condition is admittedly more opaque but no less necessary. It has been known since [38] that the prescence of a double fibre can lead to a dearth of rational points. In the aforementioned paper, the authors show that the existence of 5 double fibres implies that the rational points on the total space are not Zariski dense. In particular, they lie in a union of finitely many of the fibres. Moreover in [77, Theorem 1.4], it is shown that if a fibration has 6 double fibres then there are only finitely many points on the base whose fibre is everywhere locally soluble. Therefore excluding such fibrations is inescapable if we seek a lower bound of the order $B^{n+1}/(\log B)^A$ for the number of everywhere locally soluble fibres.

Remark. The connection between fibres whose irreducible components have high multiplicity and the density of fibres which admit a rational point is best described using the formalism of Campana points on orbifolds. Suppose that X is a smooth projective variety admitting a dominant map $\pi : Y \to X$ with geometrically integral generic fibre. A *Campana fibre* is a point $P \in X(k)$ such that $\pi^{-1}(P)$ has no irreducible component of multiplicity one. Let

$$D_\pi = \sum_{P \in X(k) \text{ Campana}} \left(1 - \frac{1}{m_P}\right) P,$$

where m_P is the minimum multiplicity of an irreducible component of $\pi^{-1}(P)$. Then the points on the base whose fibres contain a rational point are contained within the Campana points associated to this divisor, i.e.

$$\pi(Y(k)) \subset (\mathcal{X}, \mathcal{D})(\mathcal{O}_{k,S}),$$

where $\mathcal{X}, \mathcal{D}$ are integral models for X and D_π, respectively, which are regular and smooth in $\mathbb{Z}_\nu$ outside of a finite set of places S (which contains the infinite place). Since their appearance in this thesis is restricted solely to this remark, the definition of these Campana points is omitted. The intrigued and uninitiated reader should consult [90].

Proven cases of the Loughran–Smeets conjecture are very scarce in the literature. The earliest examples are due to Hooley [65] and Guo [53] for the family of diagonal ternary quadratic forms fibred over $\mathbb{P}^2$. In their original paper Loughran–Smeets proved the conjecture in the case $\Delta(\pi) = 0$ thus extending Poonen and Voloch's result (1.3.1). For conic bundles whose rank (the sum of the degrees of the non-split fibres) is at most 3, Sofos [113] has established the frequency with which the fibres are isotropic. Finally, Loughran–Matthiesen [77] have shown (in the $n = 1$ case) that the conjecture is true if the non pseudo-split fibres of π are all defined over $\mathbb{Q}$.

Rarer still are examples of problems of this type where $\Delta(\pi) > 0$ for which an asymptotic formula can be proven. There is the result of Loughran [76] which

essentially initiated this whole field of study. He proved an asymptotic formula for a Severi–Brauer variety fibred over a base which is a toric variety. Similar work of Loughran–Takloo-Bighash–Tanimoto [79] deals with Severi–Brauer varieties fibred over a wonderful compactification of an adjoint semisimple algebraic group. Further there is the work of Sofos–Visse [114] for a particularly conic bundle fibred over a hypersurface. In Chapter 6, we establish an asymptotic formula for the fibration defined by

$$\begin{aligned} \left\{\sum_{i=0}^{2} a_i X_i^2 = 0\right\} &\to \mathbb{P}^2 \\ (a_0 : a_1 : a_2; X_0 : X_1 : X_2) &\mapsto (a_0 : a_1 : a_2). \end{aligned} \tag{1.3.2}$$

This generalises the work of Serre, Hooley and Guo and can be interpreted as describing how frequently a random ternary diagonal quadratic form is isotropic.

Theorem C (with Eftymios Sofos)**.** *With H as described above and considering the family* (1.3.2)*, as $B \to \infty$, we have*

$$N_{loc}(\mathbb{P}^2, B, \pi) = \frac{7}{\Gamma(1/2)^3} \frac{B^3}{(\log B)^{\frac{3}{2}}} \prod_p \left(1 - \frac{1}{p}\right)^{\frac{3}{2}} c_p + O\left(\frac{B^3 (\log\log B)^{\frac{5}{2}}}{(\log B)^{\frac{5}{2}}}\right),$$

where c_p is the proportion of ternary conics in $\mathbb{Q}_p$ with a nontrivial $\mathbb{Q}_p$ solution.

Another particularly interesting class of varieties for which to study local-global principles are those defined by equations in which a polynomial equals the evaluation of the norm coming from a field extension.

Example 1.3.5 (The Hasse norm theorem)**.** Let K/k a cyclic extension of number fields. Suppose $\omega_1, \ldots, \omega_d$ is a basis for K as a k vector space and let $N_{K/k}$ denote the usual field norm. For $a \in k^*$, the equation

$$X_a : N_{K/k}(x_1\omega_1 + \cdots + x_d\omega_d) = a$$

satisfies the Hasse principle. Moreover weak approximation holds for X_1.

If the Hasse principle holds for X_a for all $a \in k^\times$ then we say the Hasse Norm Principle holds for k. In the example above it was proven by Hasse in [58]. It is known to fail for some non-cyclic extensions (c.f. Example 3.1.1). The question of whether or not the Hasse Norm Principle holds can be reformulated in the language of famillies of varieties. Consider

$$\begin{aligned} \mathbb{A}^{n+1} \supset V_k : \left\{N_{k/\mathbb{Q}}(x_1\omega_1 + \cdots + x_n\omega_n) = a\right\} &\to \mathbb{A}^1 \\ (x_1, \ldots, x_n; a) &\mapsto a. \end{aligned}$$

The Hasse Norm Principle holds for k exactly when every fibre satisfies the Hasse principle. The frequency of Hasse Norm Principle failures was studied by Frei–Loughran–Newton ([44], [45]) where they showed that the set of field extensions

of bounded discriminant with fixed Galois group (under a technical condition c.f. Section 3.1) satisfying Hasse norm principle has density 1. In particular, this holds for the simplest non-cyclic abelian case where the Galois group is of the form $\mathbb{Z}/2\mathbb{Z}\times\mathbb{Z}/2\mathbb{Z}$. However their result was non-explicit and hence gave no information about the asymptotic behaviour of the counting function for fields failing the Hasse norm principle. We provide an asymptotic for the number of biquadratic fields of bounded discriminant failing the Hasse norm principle, this is the main result of Chapter 3.

Theorem D. *The number of biquadratic fields $\mathbb{Q}(\sqrt{a},\sqrt{b})/\mathbb{Q}$ with discriminant bounded by X failing the Hasse norm principle is asymptotic to*

$$\frac{1}{3\sqrt{2\pi}}\sqrt{X}\sqrt{\log X}\prod_{p}\left(1+\frac{3}{2p}\right)\left(1-\frac{1}{p}\right)^{\frac{3}{2}},$$

as X goes to infinity.

More generally one could consider varieties of the form $N_{K/k}(x_1\omega_1+\cdots+x_d\omega_d)=P(t)$ for some polynomial P. In general it is possible for the Hasse principle and weak approximation to fail for varieties of this form, even when $P(t)$ is non-constant.

Example 1.3.6 ([33, Eq. 8.2])**.** This example is due to Coray. Let θ be a root of $\theta^3-7\theta^2+14\theta-7=0$. Consider the variety V defined by

$$N_{\mathbb{Q}(\theta)/\mathbb{Q}}(x+\theta y+\theta^2 y)=(t+1)(t+2).$$

This variety has rational points (e.g. $(x,y,z,t)=(0,2,0,6)$) however does not satisfy weak approximation. Indeed the rational points are not dense in the 7-adic points. To see this observe that for every point $(x,y,z,t)\in X(\mathbb{Q})$, the factor $(t+1)$ is a local norm at all places except potentially 7. At all such places ν, the cubic extension $\mathbb{Q}_\nu(\theta)/\mathbb{Q}_\nu$ of local fields is unramified and hence an element of $\mathbb{Q}_\nu$ is a local norm if and only if the valuation is divisible by 3. This clearly holds because $(t+1)(t+2)$ is a norm and thus has valuation divisible by 3, but $(t+1)$ and $(t+2)$ are coprime and thus either the valuation of $(t+1)$ is 0 or some non-zero multiple of 3. Therefore by reciprocity, it must be the case that $(t+1)$ is also a local norm at the ramified place 7. By the continuity of the map $(x_7,y_7,z_7,t_7)\mapsto t_7+1$ any point of $V(\mathbb{Q}_7)$ which is in the closure of $V(\mathbb{Q})$ must have that t_7+1 is a local norm. It suffices to find any point in $V(\mathbb{Q}_7)$ for which t_7+1 is not a local norm. Set $t_7=1$, then $(t_7+1)(t_7+2)=6$ is a local norm and $t_7+1=2$ is not.

Colliot-Thélène's conjecture predicts that such failures are explained by the Brauer–Manin obstruction. Conditional on the so-called "split polynomials conjecture", this is known to be true by work of Harpaz–Wittenberg [57, Conjecture 9.1]. However unconditional examples are limited.

Example 1.3.7. Let $k = \mathbb{Q}$ and $K = \mathbb{Q}(\sqrt{-d})$. Suppose that $P(t) \in \mathbb{Q}[t]$ is a seperable polynomial of degree 3 or 4. Then the equation

$$N_{K/k}(x + \sqrt{-d}y) = P(t),$$

defines a Châtelet surface (recall Example 7). In this case the Brauer–Manin obstruction is known to be the only one [35]-[36].

Other examples are given by :

- Colliot-Thélène–Salberger [33] for certain singular cubics corresponding to $[K : k] = 3$ and $\deg(P) \leqslant 3$,
- Schindler–Skorobogatov [101] for arbitrary extensions K/k when P splits over k and the linear factors satisfy a certain independence condition (generalising work of Heath-Brown–Skorobogatov [62], Colliot-Thélène–Harari–Skorobogatov[32] and Swarbrick-Jones [115]),
- Browning–Heath-Brown [18] when $k = \mathbb{Q}$, $[K : k] = 4$ and $P(t)$ is an irreducible quadratic, and Derenthal–Smeets–Wei[42] for k an arbitrary number field,
- Browning–Matthiesen [20] in the case that $k = \mathbb{Q}$ and P splits completely over $\mathbb{Q}$.

The frequency of Hasse principle failures for Châtelet surfaces was studied by la Bréteche–Browning [13], wherein they looked at the family

$$\begin{aligned} V : \left\{0 \neq X^2 + Y^2 = (aT^2 + b)(cT^2 + d) : ad - bc \neq 0\right\} &\to \mathbb{P}^3 \\ (X : Y : T; a : b : c : d) &\mapsto (a : b : c : d). \end{aligned} \tag{1.3.3}$$

They showed that 0% of Châtelet surfaces in this family fail the Hasse principle. In Chapter 4, we investigate what happens when the base of this fibration is changed. In particular we demonstrate that replacing $\mathbb{P}^3$ by an affine quadric produces a fibration in which a positive proportion of the fibers fail the Hasse principle.

Theorem E. *When ordering the coefficients a, b, c, d by height, if $ad - bc = \pm 1$ then a positive proportion ($\approx 23.7\ldots\%$) of the Châtelet surfaces defined by*

$$X^2 + Y^2 = (aT^2 + b)(cT^2 + d)$$

fail the Hasse principle.

The contents of this thesis represent several distinct investigations into the way that local-global principles behave in families of varieties. As such, each chapter is essentially self contained and can be read in any order. The only exception to this rule is the following chapter in which we explain some of the techniques for studying the local solubility of certain equations and those for investigating the Brauer–Manin obstruction in families. These techniques will be implemented in Chapters 3, 6 and 7 and so it is recommended that Chapter 2 be read before embarking on those chapters.

Chapter 2

Some technical results

The purpose of the present chapter is to gather together certain techniques that will feature in our investigation of local-global principles in families. The first such technique is known as the fibration method. This is a sophisticated algebro-geometric version of the Implicit Function Theorem used to study weak approximation, the Hasse principle and the Brauer–Manin obstruction in families. In the second section, we recall the definition of the Hilbert symbol and gather together lemmata which will be useful in many of the problems in later chapters.

2.1 The fibration method

We start with the proof that weak approximation holds for quadrics.

Theorem 2.1.1. *Let $Q \subset \mathbb{P}^N_k$ a smooth quadric over a number field k with local points everywhere. Then Q satisfies weak approximation.*

Proof. Let S a finite set of places and $(P_\nu)_{\nu \in S}$ a collection of local points. We aim to find a $P \in Q(k)$ which is arbitrarily ν-adically close to each of the P_ν for $\nu \in S$. By virtue of the Hasse–Minkowski theorem we know that Q has a rational point. The projection from this point to $\mathbb{P}^{N-1}$ defines a birational map $\pi : Q \dashrightarrow \mathbb{P}^{N-1}$. Possibly replacing P_ν by some point arbitrarily close to it we denote by $(P'_\nu)_\nu$ the images of the P_ν in $\mathbb{P}^{N-1}(k_\nu)$. Since weak approximation holds for $\mathbb{P}^{N-1}$ we may find a rational point P' which is arbitrarily close to each of the P'_ν and which lies in the image of π. Using the inverse of π completes the proof. □

In this proof, we have found a fibration of our total space over a base which satisfies weak approximation. Using this and nice continuity properties of the fibration, or approximation properties of the fibres, we were able to deduce the approximation property on the total space. This idea can be extended fairly simply, with the help of the Implicit Function Theorem, to the following (essentially [35, Lemma 3.9]).

Theorem 2.1.2. *Let k a number field and Y/k a smooth geometrically integral projective variety satisfying weak approximation and $Y(k) \neq \varnothing$. Let $Z \xrightarrow{\pi} Y$ a smooth fibration where each fibre has rational points and satisfies weak approximation (e.g. quadrics of dimension at least 3). Then Z satisfies weak approximation.*

Proof. Let S a finite set of places of k and $P_\nu \in Z(k_\nu)$ for each $\nu \in S$. Denote by P'_ν the image $\pi(P_\nu)$. By the Implicit Function Theorem, there exists a neighbourhood Ω_ν of P'_ν and a continuous section $s_\nu : \Omega_\nu \to Z(k_\nu)$. By weak approximation on Y, we can find a rational point P' in $Y(k)$ which lies in Ω_ν for each $\nu \in S$. The fibre $\pi^{-1}(P')$ contains points $Q_\nu = s_\nu(\Omega_\nu) \cap \pi^{-1}(P')(k_\nu)$ for $\nu \in S$ which means (possibly by shrinking Ω_ν) that Q_ν is arbitrarily close to P_ν. Now, by the conditions we placed on the fibres, we can deduce that there exists a rational point $P \in \pi^{-1}(P')(k)$ which is arbitrarily close to each Q_ν and hence to each P_ν. □

Here we've used the very strong assumption that the rational points on *every* fibre are dense in the adelic points. There has been a great deal of study over the past 40 years into how far one can relax these assumptions. Below is an example of what can be achieved if one allows simply that "most" fibres satisfy the property that the rational points are dense in the Brauer–Manin set.

Theorem 2.1.3 ([54, Théorème 4.3.1])**.** *Let V and B be geometrically integral varieties such that B satisfies weak approximation and there exists a dominant morphism $V \xrightarrow{\pi} B$ with a section s. The Brauer–Manin obstruction is the only one for any smooth projective of model of V if the following are satisfied:*

1. *The generic fibre V_η is geometrically integral, geometrically rational and s defines a smooth point in V_η,*
2. *For any smooth projective model X of V_η, $Br(\overline{X})$ is trivial and $Pic(\overline{X})$ has no torsion,*
3. *There exists a non-empty open $U \subset B$ such that $\forall b \in U$ the Brauer–Manin obstruction is the only one for all smooth proper models of V_b.*

Here we have traded off the strong assumptions on the fibres for a slightly weaker conclusion. Indeed we can no longer deduce weak approximation for that total space but rather that the only obstruction to weak approximation is the Brauer–Manin obstruction. Still in many cases this is enough as the Brauer–Manin obstruction can be computed and we can deduce a full picture of weak approximation on the total space.

2.2 Solubility of quadratic forms

2.2.1 The Hilbert symbol

Let k a local field. Inherent in many of the later chapters of this thesis are problems about the solubility of certain quadratic forms. As we saw in the

proof of the Hasse–Minkowski theorem, a key algebraic tool for the study of such problems is the Hilbert symbol. The facts collected here can be found for example in [105, Chapters 3 & 4].

Lemma 2.2.1. *For any $a, b, c \in k^\times$, we have*

(i) $\left(\frac{a,b}{k}\right) = \left(\frac{b,a}{k}\right)$.

(ii) $\left(\frac{a,c^2}{k}\right) = +1$.

(iii) $\left(\frac{a,b}{k}\right)\left(\frac{c,b}{k}\right) = \left(\frac{ac,b}{k}\right)$.

In particular suppose that k is a completion of $\mathbb{Q}$. The famous Hilbert product formula relates the Hilbert symbols over all completions of $\mathbb{Q}$. It states that if $a, b \in \mathbb{Q}^\times$ we have

$$\prod_{\nu \in \Omega_{\mathbb{Q}}} \left(\frac{a,b}{\mathbb{Q}_\nu}\right) = +1. \tag{2.2.1}$$

Furthermore there are explicit descriptions for these local Hilbert symbols.

Lemma 2.2.2. *Suppose $\nu \in \Omega_{\mathbb{Q}}$ and $a, b \in \mathbb{Q}_\nu^\times$. Then we have:*

- *If $\nu = \infty$ then $\left(\frac{a,b}{\mathbb{Q}_\nu}\right) = -1$ if and only if both a and $b < 0$.*
- *If $\nu = p$ an odd prime, then*

$$\left(\frac{a,b}{\mathbb{Q}_\nu}\right) = \left(\frac{-1}{p}\right)^{v_p(a)v_p(b)} \left(\frac{a/p^{v_p(a)}}{p}\right)^{v_p(b)} \left(\frac{b/p^{v_p(b)}}{p}\right)^{v_p(a)}.$$

- *If $\nu = 2$ then write $a = 2^\alpha u, b = 2^\beta v$ where $u, v \in \mathbb{Z}_2^\times$. Then $\left(\frac{a,b}{\mathbb{Q}_\nu}\right) = (-1)^{\frac{u-1}{2}\frac{v-1}{2} + \alpha\frac{v^2-1}{8} + \beta\frac{u^2-1}{8}}$.*

Now suppose k is the completion of a number field at some place. Suppose $f = a_1X_1^2 + \ldots a_nX_n^2$ where $a_i \in k$ then we introduce the invariants

1. $d(f) = a_1 \cdot \ldots a_n$ in $k^\times/(k^\times)^2$,
2. $\epsilon(f) = \prod_{i<j} \left(\frac{a_i,a_j}{k}\right)$.

Using these we can understand when the form f has a non-trivial solution.

Theorem 2.2.3. *For f to represent 0 it is necessary and sufficient that one of the following hold:*

1. *$n = 2$ and $d = -1$,*
2. *$n = 3$ and $\left(\frac{-1,-d}{k}\right) = \epsilon$,*
3. *$n = 4$ and either $d \neq 1$ or $d = 1$ and $\epsilon = \left(\frac{-1,-1}{k}\right)$,*
4. *$n \geqslant 5$.*

Combining this with the Hasse–Minkowski theorem gives a complete desciption of the solubility of diagonal quadratic forms over number fields.

2.2.2 Analytic techniques

In [47], Friedlander–Iwaniec studied ternary quadratic forms defined by equations

$$Z^2 = aX^2 + bY^2,$$

where a, b were assumed for simplicity to be squarefree and coprime. This equation has a solution exactly when all the local Hilbert symbols satisfy $\left(\frac{a,b}{\mathbb{Q}_\nu}\right) = +1$. Using the description in Lemma 2.2.2, we can write down an indicator function for this property. Evidently if $p \nmid ab$ then $\left(\frac{a,b}{\mathbb{Q}_p}\right) = +1$. Suppose $p \mid a$ then

$$\mathbf{1}_{\left(\frac{a,b}{\mathbb{Q}_p}\right)=+1} = \frac{1}{2}\left(1 + \left(\frac{b}{p}\right)\right).$$

Note here how we have used the fact that $\mu^2(ab) = 1$. This simplifies the Friedlander–Iwaniec counting problem to evaluating

$$\sum_{\substack{a\leqslant A\\ b\leqslant B}} \frac{\mu^2(ab)}{\tau(ab)} \prod_{p\mid ab}\left(1+\left(\frac{a}{p}\right)\right)\left(1+\left(\frac{b}{p}\right)\right) = \sum_{\substack{k\ell\leqslant A\\ mn\leqslant B}} \frac{\mu^2(k\ell mn)}{\tau(k\ell mn)}\left(\frac{k\ell}{m}\right)\left(\frac{mn}{k}\right). \tag{2.2.2}$$

Restricting for the moment to the case when $a \equiv b \equiv 1 \bmod 4$ then we can use quadratic reciprocity to reduce this sum to

$$\sum_{\substack{k\ell\leqslant A\\ mn\leqslant B\\ k\equiv \ell \bmod 4\\ m\equiv n \bmod 4}} \frac{\mu^2(k\ell mn)}{\tau(k\ell mn)}\left(\frac{\ell}{m}\right)\left(\frac{n}{k}\right). \tag{2.2.3}$$

To attack these character sums they employed the following lemma.

Lemma 2.2.4 ([47, Corollary 2])**.** *Let $q = q_1q_2 \in \mathbb{N}$ where $(q_1, q_2) = 1$, $(ad, q) = 1$ and χ is a Dirichlet character modulo q_2. Then for any $C > 0$ we have*

$$\sum_{\substack{n\leqslant x\\ (n,d)=1\\ n\equiv a \bmod q_1}} \mu^2(n)\frac{\chi(n)}{\tau(n)} = \delta_\chi \frac{c_0F(dq)}{\phi(q_1)}\frac{x}{\sqrt{\log x}}\left\{1 + O\left(\frac{(\log\log x)^{\frac{3}{2}}}{\log x}\right)\right\}$$
$$+ O_C\left(\tau(d)qx(\log x)^{-C}\right),$$

where δ_χ is 1 if χ is principal and 0 otherwise, and

$$c_0 = \pi^{-\frac{1}{2}}\prod_p\left(1+\frac{1}{2p}\right)\left(1-\frac{1}{p}\right)^{\frac{1}{2}}$$
$$F(r) = \prod_{p\mid r}\left(1+\frac{1}{2p}\right)^{-1}.$$

Since the application of these lemmas forms a large part of various proofs throughout the rest of the thesis, we have elected to provide sketchs of their proofs. It is hoped that this will make the thesis more self-contained.

Sketch. The lemma is proven by an application of the Truncated Perron Formula applied to the associated Dirichlet series

$$Z_d(s,\chi) := \sum_{\gcd(n,d)=1} \frac{\mu^2(n)\chi(n)}{\tau(n)n^s}.$$

The initial line segment of integration is set to be $[2-iT, 2+iT]$ and then moved back to $[1-\eta-iT, 1-\eta+iT]$ via a keyhole contour, where $T = \exp(c\sqrt{\log x})$ and $q = \frac{c(\epsilon)}{q^\epsilon \log T}$. The contributions from the vertical and horizontal line segments (not on the real axis) are dealt with in a classical manner. We will only record here a proof of the main term as it is the most non-standard part of the proof. When χ is not principal, we may write

$$Z_d(s,\chi) = L_{dq}(s)\zeta(s)^{\frac{1}{2}}R(s),$$

where

$$L_{dq}(s) = \prod_{p|dq}\left(1+\frac{1}{2p^s}\right)^{-1}$$
$$R(s) = \prod_{p}\left(1+\frac{1}{2p^s}\right)\left(1-\frac{1}{p^s}\right)^{\frac{1}{2}}.$$

When $s = \sigma \pm \epsilon i$, for small ϵ and $\sigma \in (\frac{2}{3}, 1]$, we use the approximations

$$\zeta(s)^{\frac{1}{2}} = \frac{\mp i}{\sqrt{1-\sigma}}\left(1+O(1-\sigma)\right),$$
$$L_{dq}(s) = \sum_{\substack{h|d^\infty \\ h\leqslant\sqrt{x}}}\frac{a_h}{h^s} + O\left(\frac{\tau(dq)^{\frac{1}{4}}}{x}\right),$$
$$R(s) = R(1)\left(1+O(1)\right).$$

This means the contribution to the contour from the real line segments is given by

$$\frac{R(1)}{\pi}\int_{1-\eta}^{1}\frac{x^\sigma}{\sqrt{1-\sigma}}\left(\sum_{\substack{h|d^\infty \\ h\leqslant\sqrt{x}}}\frac{a_h}{h^\sigma}\right) + O\left(x^{\frac{3}{4}}\tau(dq)\right) + O\left(\int_{1-\eta}^{1}x^\sigma\sqrt{1-\sigma}\right).$$

The main contribution is handled by bounding the integral by

$$\int_{-\infty}^{1}\frac{y^\sigma}{\sqrt{1-\sigma}}\mathrm{d}\sigma = 2y\int_0^\infty e^{-z^2\log y}\mathrm{d}z = \frac{\sqrt{\pi}y}{\sqrt{\log y}},$$

with $y = x/h$. This yields

$$\frac{R(1)x}{\sqrt{\pi}\sqrt{\log x}} \sum_{\substack{h|(dq)^\infty \\ h\leqslant\sqrt{x}}} \frac{a_h}{h}\left(1 + O\left(\frac{\log h}{\log x}\right)\right).$$

Extending the sum over h and using the bounds

$$\sum_{h|(dq)^\infty} \frac{|a(h)|}{h}\log h \ll \prod_{p|dq}\left(1-\frac{1}{2p}\right)^{-1}\sum_{p|dq}\frac{\log p}{p} \ll (\log\log 3dq)^{\frac{3}{2}},$$

completes the proof. □

One obstacle to applying this lemma to all the sums of the form (2.2.3) is the quality of the error term, typical for such a Siegel–Walfisz type result. Namely when the modulus of the character is not smaller than a log power of the length of the sum (e.g. when $q \approx x^\theta$, for some small $\theta > 0$) then the error term dominates the main term. To handle the contribution arising when the variables k, ℓ, m and n are possibly large, Friedlander–Iwaniec appealed to the celebrated large sieve for quadratic characters.

Lemma 2.2.5 ([61, Corollary 4]). *Let a_m, b_n complex sequences of modulus bounded by one, supported on odd square-free integers $m \leqslant M$ and $n \leqslant N$. Then for any $\epsilon > 0$, we have*

$$\sum_{m\leqslant M}\sum_{n\leqslant N} a_m b_n \left(\frac{m}{n}\right) \ll_\epsilon (MN)^{1+\epsilon}\left(M^{-\frac{1}{2}} + N^{-\frac{1}{2}}\right).$$

On occasion we will have need of an effective version of this result, for which we quote Friedlander–Iwaniec.

Lemma 2.2.6 ([47, Lemma 2]). *Let a_m, b_n as above. Then we have*

$$\sum_{m\leqslant M}\sum_{n\leqslant N} a_m b_n \left(\frac{m}{n}\right) \ll \left(MN^{\frac{5}{6}} + M^{\frac{5}{6}}N\right)(\log 3MN)^{\frac{7}{6}},$$

where the implied constant is absolute.

Proof. Let $B(M,N)$ be the sum which we aim to estimate and assume that $N \leqslant M$. Applying Hölder's inequalities yields

$$|B(M,N)|^3 \leqslant \left(\sum_{n\leqslant N}|b_n|^{3/2}\right)^2 \sum_{n\leqslant N}\left|\sum_{m\leqslant M} a_m\left(\frac{m}{n}\right)\right|^3$$

Expanding the cube and re-arranging the sums we get the bound

$$|B(M,N)|^3 \leqslant N^2 \sum_{\ell\leqslant M^3}\tau_3(\ell)\left|\sum_{n\leqslant N}\gamma_n\left(\frac{\ell}{n}\right)\right|,$$

where τ_3 denotes the 3-fold divisor function. Now applying the Cauchy–Schwarz inequality, we get

$$\begin{aligned}|B(M,N)|^6 &\leqslant N^4 \sum_{\ell \leqslant M^3} \tau_3(\ell)^2 \sum_{\ell \leqslant M^3} \left| \sum_{n \leqslant N} \gamma_n \left(\frac{\ell}{n}\right) \right|^2 \\ &\ll N^4 M^6 (\log 2M)^6 \sum_{n_1 \leqslant N} \sum_{n_2 \leqslant N} \left| \sum_{\ell \leqslant M^3} \left(\frac{\ell}{n_1 n_2}\right) \right|.\end{aligned}$$

The main contribution to the remaining sum occurs when $n_1 n_2$ is a square which occurs for $\ll N \log N$ choices of n_1, n_2 and otherwise the sum over ℓ is bounded by $n_1 n_2$ giving a bound of $M^3 N \log N + N^4$ which completes the proof. □

This approach to measuring how frequently fibers in a family of quadratic forms are soluble will form the basis for our approach to the related, more complicated settings faced in Chapters 3, 6 and 7.

Chapter 3

Hasse norm principle failures in biquadratics

3.1 Introduction

The contents of this chapter are primarily based on [99], with some slight improvements. Let K/k be an extension of number fields. There exists a map $N_{K/k} : K^* \to k^*$ known as the field norm defined as follows: Let $\{x_1, \ldots, x_n\}$ be a basis for K as a $k-$vector space and for any $a \in K^*$ denote by ϕ_a the linear map $x \mapsto xa$, then $N_{K/k}(a)$ is the determinant of the matrix associated to ϕ_a. An element of k which is in the image of the norm is said to be a *global norm* of K. This map naturally extends to local extensions K_μ/k_ν, where μ and ν are places of K and k, respectively. Analogously, elements of k_ν in the image of this map are referred to as *local norms*. We may consider the norm (abusing notation) to be defined on the invertible adeles $N_{K/k} : \mathbb{A}_K^* \to \mathbb{A}_k^*$.

Remark. The invertible adeles of K are also known as the idéles and denoted $\mathbb{I}_K$, however they are usually endowed with a topology other than the one inherited as a subspace of $\mathbb{A}_K$ (namely the restricted product topology with respect to the unit groups $\mathcal{O}_{K_\nu}$, c.f. Chapter 1 p. 7).

We say that the Hasse norm principle (frequently just HNP) holds if

$$k^* \cap N_{K/k}\mathbb{A}_K^* = N_{K/k}K^*,$$

i.e. every element of k which is a local norm everywhere is also a global norm. The first example of a class of extensions for which this property always holds was discovered by Hasse [58], who showed that the HNP always holds if K/k is a cyclic extension (c.f. Chapter 1, Example 1.3.5). Since then many more such classes have been discovered:

1. $[K : k]$ prime [5, Section 2, Lemma 4],
2. $[K : k] = n$ and $\mathrm{Gal}(K^{\mathrm{norm}}/K) \cong D_n$ [6, Satz 1],

3. $[K:k]=n$ and $\mathrm{Gal}(K^{\mathrm{norm}}/K)\cong S_n$ [122],

4. $[K:k]=n$ and $\mathrm{Gal}(K^{\mathrm{norm}}/K)\cong A_n$ for $n\geqslant 5$ [81].

(Note: K^{norm} here denotes the normal closure of K/k.) However even in the simplest non-cyclic case HNP can fail.

Example 3.1.1 ([24, Exercise 5.3])**.** Every rational square in the biquadratic field $\mathbb{Q}(\sqrt{13},\sqrt{17})$ is a local norm everywhere [45, Lemma 4.4] but 5^2, for example, is not a global norm.

There is a geometric interpretation for these failures of the local-global principle. The defect of the Hasse norm principle is measured by the *knot group*, $\text{Ш}(K/k)=\left(k^*\cap N_{K/k}\mathbb{A}_K^*\right)/N_{K/k}K^*$ which is isomorphic to the Tate–Shafarevich group $\text{Ш}(T)$ of the associated norm one torus $T=R^1_{K/k}\mathbb{G}_m$. For fixed $c\in k^*$, the affine variety defined by the equation $N_{K/k}(x)=c$ is a principal homogeneous space for T. Therefore another way of thinking of the Hasse norm principle is as the statement that all principal homogeneous spaces for the norm one torus satisfy the Hasse principle. One may also ask about weak approximation for T. In analogy with the knot group, we define the defect of weak approximation as $A(T)=\prod_\nu T(k_\nu)/\overline{T(k)}$. There exists a short exact sequence due to Voskresenskiĭ [121, Theorem 6] connecting these two groups:

$$0\to A(T)\to \mathrm{H}^3(G,\mathbb{Z})^\sim\to \text{Ш}(T)\to 0. \tag{3.1.1}$$

Here $A^\sim$ denotes the dual $\mathrm{Hom}(A,\mathbb{Q}/\mathbb{Z})$. Weak approximation can also fail for the norm one tori associated to non-cyclic extensions. Hence it is natural to ask, for Galois groups for which HNP (or weak approximation) is not guaranteed, how frequently do failures occur?

For abelian extensions, this question was addressed in a pair of papers by Frei–Loughran–Newton [44, 45]. For a fixed number field k, they showed that for every non-cyclic abelian group G there exists an extension K/k with $\mathrm{Gal}(K/k)\cong G$ such that HNP fails. Moreover they proved that when ordered by conductor 0% of extensions with a fixed non-cyclic abelian Galois group fail HNP. The situation when ordering the fields by discriminant is slightly more complicated. Suppose that $n,r\in\mathbb{Z}_{\geqslant 0}$, Q is the smallest prime dividing n and $G\cong\mathbb{Z}/n\mathbb{Z}\oplus(\mathbb{Z}/Q\mathbb{Z})^r$, then 0% of extensions K/k with Galois group G fail HNP. When the Galois group is not of this form then a positive proportion of extensions fail HNP. This is an interesting example of a phenomenon discussed by Matchett-Wood [85], where she makes the case that ordering fields by conductor is a more natural choice (in her terms, the conductor is a "fair counting function" and the discriminant is not). Frei–Loughran–Newton also proved that a positive proportion of norm one tori associated to extensions with abelian, non-cyclic Galois group G fail weak approximation. These results are strictly density statements and the proofs do not give a precise estimate for the frequency of extensions failing HNP. The purpose of this chapter is to study the simplest abelian non-cyclic case where $k=\mathbb{Q}$ and $G\cong(\mathbb{Z}/2\mathbb{Z})^2$. Let Δ_K denote the absolute discriminant of the biquadratic field $K=\mathbb{Q}(\sqrt{a},\sqrt{b})$. We

will develop an asymptotic formula for the number of biquadratic extensions of bounded discriminant which fail the Hasse norm principle. However our first result, for the purpose of comparison, is a proof of the total number of biquadratic fields of bounded discriminant.

Theorem 3.1.2. *There exists a polynomial P of degree 2 with leading coefficient*

$$\frac{23}{960}\prod_p \left(1-\frac{1}{p}\right)^3\left(1+\frac{3}{p}\right)$$

such that the number of biquadratic fields $K/\mathbb{Q}$ with discriminant bounded by X is

$$\sqrt{X}P(\log X)+O\left(X^{\frac{1}{4}}\log^{\frac{13}{4}} X\right).$$

Remark. The coefficients of the lower order terms in P are explicitly computable from the proof, however we have chosen not to only write out the main term here.

This result is not entirely new. The main term was originally proven by Baily [2, Theorem 8] in 1980 and also follows from work of Wright [124, Theorem 1.2]. However neither of these results give strong information about the error term (second order terms were first shown by Cohen–Diaz-y-Diaz–Olivier [26, Section 2.5]). The proof of Theorem 3.1.2 is included as it illustrates the method of proof for Theorem 3.1.3, which is the main result of this chapter. Recently la Bréteche–Kurlberg–Shparlinski [12] have extended the method of proof of Theorem 3.1.2 to arbitrary multiquadratic field extensions and further they substantially improved upon the original error term given in [99]. The same problem was also independently addressed by Fritsch [48]. We incorporate some of the ideas of [12] into the present proof of Theorem 3.1.2. The main result of this section is the following asymptotic formula for the fields of bounded discriminant which fail the Hasse norm principle.

Theorem 3.1.3. *The number of biquadratic fields $K/\mathbb{Q}$ with discriminant bounded by X and such that the Hasse norm principle fails is*

$$\frac{1}{3\sqrt{2\pi}}\sqrt{X\log X}\prod_p\left(1-\frac{1}{p}\right)^{\frac{3}{2}}\left(1+\frac{3}{2p}\right)+O\left(\sqrt{X}\right).$$

This result together with Theorem 3.1.2 recovers the 0% density in the Frei–Loughran–Newton result [44, Theorem 1.1]. Furthermore, we include here as a simple corollary of the techniques used to prove Theorems 3.1.2 and 3.1.3 the associated asymptotic formulae for the situation in which the fields are ordered by conductor.

Theorem 3.1.4. *The number of biquadratic fields of conductor bounded by X is*

$$\frac{78}{23}XP(\log X)+O(X^{\frac{1}{2}}\log^{\frac{13}{4}} X),$$

with P the same as in Theorem 3.1.2. The number of these fields which fail the Hasse norm principle is

$$\frac{2}{\sqrt{\pi}} X \sqrt{\log X} \prod_p \left(1 - \frac{1}{p}\right)^{\frac{3}{2}} \left(1 + \frac{3}{2p}\right) + O(X).$$

An orthogonal investigation to ours and that of Frei–Loughran–Newton, which is more in the spirit of the fibration questions discussed in Chapter 1, was conducted by Browning–Newton [21]. They fix an extension $K/\mathbb{Q}$ and investigate the frequency with which elements of $\mathbb{Q}$ are representable as everywhere local norms but not global norms. They demonstrated that in fact the proportion of everywhere local norms which are not global is given by $1 - \frac{1}{\#\text{Ш}(K/\mathbb{Q})}$. We also take this opportunity to point out that some results concerning failures of the Hasse norm principle in non-abelian settings have been established by Macedo [81] and Macedo–Newton [82]. At present no quantitative results have been proven (the harmonic analysis approach of [44] requires the input of class field theory so breaks down in the setting of non-abelian extensions) but there is hope that the techniques commonly collectively referred to as "Bhargavology" can be used to provide asymptotics.

3.2 Criterion for HNP failure

In this section, we'll describe a criterion on the integers a and b that determine when the extension $K = \mathbb{Q}(\sqrt{a}, \sqrt{b})$ fails the Hasse norm principle. We can then sum over the a and b satisfying this criterion to get Theorem 3.1.3.

First, we describe how to count a and b that define unique biquadratic extensions of $\mathbb{Q}$. Note that K has 3 quadratic subfields

$$k_1 = \mathbb{Q}(\sqrt{a}),\ k_2 = \mathbb{Q}(\sqrt{b}) \text{ and } k_3 = \mathbb{Q}(\sqrt{ab}/(a,b)).$$

Each of these quadratic fields can be uniquely identified by a single squarefree integer so fix $k_1 = \mathbb{Q}(\sqrt{a})$ and $k_2 = \mathbb{Q}(\sqrt{b})$. Let $m = (a,b)$ so that

$$a = ma_1, b = mb_1 \text{ and } (a_1, m) = (a_1, b_1) = (b_1, m) = 1. \tag{3.2.1}$$

Then

$$k_1 = \mathbb{Q}(\sqrt{ma_1}),\ k_2 = \mathbb{Q}(\sqrt{mb_1}) \text{ and } k_3 = \mathbb{Q}(\sqrt{a_1 b_1}).$$

It is certainly true that specifying m, a_1, b_1 will determine K, so long as each of the products ma_1, mb_1 and $a_1 b_1$ are not squares. Moreover there are 6 ways to produce the same triple, by relabelling the quadratic subfields.

We can write the discriminant of K in terms of (m, a_1, b_1) as follows. By the conductor-discriminant formula (see e.g. [49, Ch. 8, 7.23]) we can express the discriminant of K, denoted Δ_K, in terms of the discriminants of its quadratic subfields by

$$\Delta_K = \Delta_{k_1} \Delta_{k_2} \Delta_{k_3}.$$

Recall that

$$\operatorname{disc}\left(\mathbb{Q}(\sqrt{d})\right) = \begin{cases} d & \text{if } d \equiv 1 \bmod 4 \\ 4d & \text{if } d \equiv 2 \text{ or } 3 \bmod 4. \end{cases} \tag{3.2.2}$$

We observe that it is not possible for just one of the integers ma_1, mb_1 and a_1b_1 to be congruent to 2 or 3 mod 4. For if $ma_1 \equiv 2 \bmod 4$ then either m or $a_1 \equiv 2 \bmod 4$ and hence so is their product with b_1. Moreover $ma_1 \equiv 3 \bmod 4 \iff m \equiv -a_1 \bmod 4$ so either $m \equiv -b_1 \bmod 4$ or $a_1 \equiv -b_1 \bmod 4$. Therefore

$$\Delta_K = c^2 m^2 a_1^2 b_1^2 \tag{3.2.3}$$

where c is either 1 if all the k_i are in the first case of (3.2.2), 4 if exactly one k_i is in the first case of (3.2.2) or 8 if all the k_i are in the second case.

We now turn our attention to how to identify Hasse norm principle failures. Tate gave an explicit description of the knot group in terms of Galois cohomology.

Theorem 3.2.1. *Let G be a finite group and K/k an extension of number fields with Galois group G. Then there is a canonical isomorphism*

$$\text{Ш}(K/k)^{\sim} \cong Ker\left(H^3(G,\mathbb{Z}) \to \prod_{\nu} H^3(G_\nu,\mathbb{Z})\right),$$

where G_ν denotes the decomposition group at the place ν of k.

Proof. See e.g. [92, Theorem 6.11]. □

From this theorem, Hasse's norm theorem quickly follows since if G is cyclic then $H^3(G,\mathbb{Z}) = 0$. In the case $G = (\mathbb{Z}/2\mathbb{Z})^2$ we have $H^3(G,\mathbb{Z})^{\sim} \cong \mathbb{Z}/2\mathbb{Z}$ (this can be proven for example by combining [44, Lemmas 6.4 and 6.5]). Therefore if all the decomposition groups are cyclic the kernel of the map above is all of $H^3(G,\mathbb{Z})$ hence the knot group is nontrivial and the Hasse norm principle fails. Conversely if any of the decomposition groups is isomorphic to $(\mathbb{Z}/2\mathbb{Z})^2$ then the kernel is empty and HNP holds. The following lemma expresses this criterion as a condition on the variables m, a_1 and b_1.

Lemma 3.2.2. *Let $(m, a_1, b_1) \equiv (\epsilon_1, \epsilon_2, \epsilon_3) \bmod 4$. Then*

1. *When $\epsilon_1 = \epsilon_2 = \epsilon_3$, K fails the Hasse norm principle if and only if, for p odd, all of the following hold:*

   (i) $p \mid a_1 \implies \left(\frac{mb_1}{p}\right) = +1,$

   (ii) $p \mid b_1 \implies \left(\frac{ma_1}{p}\right) = +1,$

   (iii) $p \mid m \implies \left(\frac{a_1b_1}{p}\right) = +1.$

2. *When $\epsilon_1 = \epsilon_2 \neq \epsilon_3$, K fails the Hasse norm principle if and only if, for p odd, all of the following hold:*

(i) $p \mid a_1 \implies \left(\frac{mb_1}{p}\right) = +1$,

(ii) $p \mid b_1 \implies \left(\frac{ma_1}{p}\right) = +1$,

(iii) $p \mid m \implies \left(\frac{a_1 b_1}{p}\right) = +1$,

(iv) $m \equiv a_1 \bmod 8$.

Similarly for $\epsilon_2 = \epsilon_3 \neq \epsilon_1$ *and* $\epsilon_3 = \epsilon_1 \neq \epsilon_2$.

3. *If* ϵ_1, ϵ_2 *and* ϵ_3 *are pairwise distinct, then* K *satisfies the Hasse norm principle.*

Proof. The Hasse norm principle fails in biquadratics if and only if all decomposition groups are cyclic. Hence to come up with a criterion for Hasse norm principle failure we need to ensure that all decomposition groups are proper subgroups of the Galois group $(\mathbb{Z}/2\mathbb{Z})^2$. Therefore we need every rational prime to split in K (i.e. the ideal $p\mathcal{O}_K$ in K can be written as the product of two or four prime ideals in $\mathcal{O}_K$).

A prime splits in K if and only if it splits in at least one of the three quadratic subfields $\mathbb{Q}(\sqrt{ma_1})$, $\mathbb{Q}(\sqrt{mb_1})$ and $\mathbb{Q}(\sqrt{a_1b_1})$.

1. Firstly it cannot be the case that $a_1 \equiv m \equiv b_1 \equiv 2 \bmod 4$ since a_1,b_1 and m are pairwise coprime, hence in this case they must all be odd. If $a_1 \equiv m \equiv b_1 \bmod 4$ then the only primes that ramify in $\mathbb{Q}(\sqrt{ma_1})$ are those dividing ma_1, therefore since m, a_1 and b_1 are pairwise coprime, no prime ramifies in all three quadratic subextensions. An odd unramified prime p splits in $\mathbb{Q}(\sqrt{a}) \iff \left(\frac{a}{p}\right) = +1$ otherwise it remains inert. This means a prime cannot be inert in all three subfields. Hence we must ensure that all primes that ramify in two of the subfields split in the third.

2. If $m \equiv a_1 \not\equiv b_1 \bmod 4$ then $ma_1 \equiv 1 \bmod 4$, $mb_1 \equiv 2$ or $3 \bmod 4$ and $a_1b_1 \equiv 2$ or $3 \bmod 4$. Note that since m and a_1 are coprime they cannot both be congruent to 2, so they must be odd. We see that the rational prime 2 also ramifies in the subfields $\mathbb{Q}(\sqrt{mb_1})$ and $\mathbb{Q}(\sqrt{a_1b_1})$. Therefore we must also ensure that 2 splits in $\mathbb{Q}(\sqrt{ma_1})$ so we must impose the extra condition $ma_1 \equiv 1 \bmod 8$.

3. In this case we have $ma_1 \equiv 2$ or $3 \bmod 4$, $mb_1 \equiv 2$ or $3 \bmod 4$ and $a_1b_1 \equiv 2$ or $3 \bmod 4$ so 2 ramifies in all 3 quadratic subextensions hence is totally ramified in K. Therefore the Hasse norm principle holds.

□

Lastly in this section we note that the conditions for weak approximation on the norm one torus are exactly the same as HNP failure in this case. Indeed when K/k is biquadratic $\mathrm{H}^3(G, \mathbb{Z})^\sim \cong \mathbb{Z}/2\mathbb{Z}$. Hence by (3.1.1) either $A(T) \cong \mathbb{Z}/2\mathbb{Z}$

and $Ш(T) = 0$ or vice versa. This means that in fact the biquadratic extensions for which HNP fails and for which weak approximation holds on the norm one torus coincide. Therefore one could view our Theorem 3.1.3 equivalently as a count of those extensions satisfying the weak approximation condition.

3.3 Initial field count

In this section we count the number of biquadratic fields of bounded discriminant. In the previous section we saw that the number of biquadratic extensions of $\mathbb{Q}$ with discriminant bounded by X is equal to

$$\frac{1}{6}\#\{(a,b) \in \mathbb{Z}^2 : \mathrm{Disc}(\mathbb{Q}(\sqrt{a},\sqrt{b})) \leqslant X \text{ and } [\mathbb{Q}(\sqrt{a},\sqrt{b}) : \mathbb{Q}] = 4\}.$$

Then we showed how each element in the set above corresponds to a triple of squarefree, pairwise coprime integers, (m_1, m_2, m_3) say, such that each of the products $m_i m_j$ is not a rational square. For non-unit m_i this is guaranteed by the squarefree and pairwise coprimality conditions, but we must further remove the cases where $m_i = m_j = \pm 1$ for $i \neq j$. Ignoring triples with this condition in the sum (3.3.2) below will incur an error of size $\frac{2}{\zeta(2)}\sqrt{X} + O(X^{1/4})$ thus not affecting the statement of Theorem 3.1.2. Let $\boldsymbol{\delta} = (\delta_2, \delta_3)$ where δ_i is the sign of m_i. Observe that the highest power of 2 dividing $m_1 m_2 m_3$ is either 0 or 1. To keep track of this we write

$$m_1 = 2^{\mu} m_1', \; m_2 = \delta_2 2^{\alpha} m_2' \text{ and } m_3 = \delta_3 2^{\beta} m_3', \tag{3.3.1}$$

where $2 \nmid m_1' m_2' m_3'$. The discriminant of the biquadratic field associated to the triple (m_1, m_2, m_3) is given by $C^2 (m_1' m_2' m_3')^2$ where

$$C = \begin{cases} 1 \text{ if } \mu = \alpha = \beta = 0 \text{ and } m_1 \equiv m_2 \equiv m_3 \bmod 4, \\ 4 \text{ if } \mu = \alpha = \beta = 0 \text{ and } \exists i \text{ such that } m_i \not\equiv m_j \bmod 4 \\ 8 \text{ if } \mu = 1 \text{ and } m_2 \equiv m_3 \bmod 4, \\ 8 \text{ if } \alpha = 1 \text{ and } m_1 \equiv m_3 \bmod 4, \\ 8 \text{ if } \beta = 1 \text{ and } m_1 \equiv m_2 \bmod 4, \\ 16 \text{ otherwise.} \end{cases}$$

Denote by $E(C)$ the set of residue classes mod 4 in which a triple (m_1, m_2, m_3) may reside so that the discriminant of the associated field features the constant C. Now the count for biquadratic extensions of bounded discriminants is given

by

$$\frac{1}{6}\sum_{C\in\{1,4,8,16\}}\sum_{\boldsymbol{\delta}\in\{\pm1\}^2}\sum_{\substack{\mu,\alpha,\beta\in\{0,1\}\\ \mu+\alpha+\beta\in\{0,1\}}}\sum_{\substack{m_1'm_2'm_3'\leqslant\sqrt{X}/C\\ 2\nmid m_1'm_2'm_3'\\ (m_1,m_2,m_3)\bmod 4\in E(C)}}\mu^2(m_1'm_2'm_3')$$

$$=\frac{1}{6}\sum_{C\in\{1,4,8,16\}}\sum_{\substack{m_1'm_2'm_3'\leqslant\sqrt{X}/C\\ 2\nmid m_1'm_2'm_3'}}\mu^2(m_1'm_2'm_3')\sum_{\boldsymbol{\delta}\in\{\pm1\}^2}\sum_{\substack{\mu,\alpha,\beta\in\{0,1\}\\ \mu+\alpha+\beta\in\{0,1\}\\ (m_1,m_2,m_3)\bmod 4\in E(C)}}1.$$

Letting $T(C)$ be the number of choices for $\mu,\alpha,\beta,\delta_2$ and δ_3, the above can be written as

$$\frac{1}{6}\sum_{C\in\{1,4,8,16\}}T(C)\sum_{\substack{m_1'm_2'm_3'\leqslant\sqrt{X}/C\\ 2\nmid m_1'm_2'm_3'}}\mu^2(m_1'm_2'm_3'). \tag{3.3.2}$$

In particular,

$$T(1)=1,T(4)=3,T(8)=6 \text{ and } T(16)=6.$$

The formulation here differs slightly from [99] and is inspired by [12, Section 6.2]. The problem now is reduced to the study of the sum

$$A(x)=\sum_{\substack{n\leqslant x\\ n \text{ odd}}}\mu^2(n)3^{\omega(n)}.$$

In [99] this sum was tackled by an appeal to the Selberg–Delange method. It turns out that the full generality of results of this type is not necessary and one can produce better error terms by directly studying the associated Dirichlet series. In [12], the authors apply Perron's formula to this Dirichlet series and use Weyl's pointwise subconvexity bound for $\zeta(\frac{1}{2}+it)$ to obtain a power saving error term. We now follow that approach but incorporate information about moments of ζ to produce an even stronger error term.

Lemma 3.3.1. *There exists a polynomial P of degree 2 such that*

$$A(x)=xP(\log x)+O\left(x^{\frac{1}{2}}\log^{\frac{13}{4}}x\right)$$

Moreover the leading coefficient of P is given by

$$\frac{1}{5}\prod_p\left(1-\frac{1}{p}\right)^3\left(1+\frac{3}{p}\right).$$

Proof. Consider the Dirichlet series

$$F(s)=\sum_{n \text{ odd}}\frac{\mu^2(n)3^{\omega(n)}}{n^s}.$$

This series is absolutely convergent for $\mathrm{Re}(s) > 1$ and in this region admits an Euler product $\prod_{p>2}\left(1+\frac{3}{p^s}\right)$. We can write $F(s) = (1+\frac{3}{2^s})^{-1}\zeta(s)^3 G(s)$ where G is an absolutely convergent Dirichlet series for $\mathrm{Re}(s) > \frac{1}{2}$. By Perron's formula [118, Lemma 3.19], we have

$$A(x) = \frac{1}{2\pi i}\int_{1+\frac{1}{\log x}-iT}^{1+\frac{1}{\log x}+iT} F(s)X^s\frac{\mathrm{d}s}{s} + O_\epsilon\left(\frac{x^{1+\epsilon}}{T}\right).$$

The only pole of $F(s)$ in the region $\mathrm{Re}(s) > \frac{1}{2}$ occurs at $s = 1$ so we may move the line of integration back to $\mathrm{Re}(s) = \frac{1}{2}$ using a keyhole contour. The integral of the circle around $s = 1$ will give $xQ(\log x)$ for Q some polynomial of degree 2 since the pole has order 3. The leading coefficient is given by the residue $\frac{1}{2!}(1+\frac{3}{2})^{-1}G(1)$. The contribution from the vertical line segment connecting $s = \frac{1}{2} - iT$ and $s = \frac{1}{2} + iT$ is bounded by

$$\int_{\frac{1}{2}-iT}^{\frac{1}{2}+iT} |F(s)|x^{\mathrm{Re}(s)}\frac{\mathrm{d}s}{|s|} \ll x^{\frac{1}{2}}\int_{-T}^{T}|\zeta(\frac{1}{2}+it)|^3\frac{\mathrm{d}t}{1+|t|}.$$

To bound this we will use an upper bound for the third moment of ζ on the half line. The $2k^{\text{th}}$ moments of ζ on the half line, denoted $M_k(T)$ are classical objects of study in analytic number theory. The current state of the art is the bound $M_k(T) \ll T\,(\log T)^{k^2}$ for any real k in $[0,2]$ [60]. Plugging this into the above yields an upper bound for the third moment of $\ll T\,(\log T)^{\frac{9}{4}}$. To recover information about the vertical integral, we break the range into dyadic intervals

$$2\sum_{\ell=0}^{\log_2 T}\int_{T/2^{\ell+1}}^{T/2^\ell}|\zeta(\frac{1}{2}+it)|^3\frac{\mathrm{d}t}{1+|t|} + \int_0^{\frac{1}{2}}|\zeta(\frac{1}{2}+it)|^3\frac{\mathrm{d}t}{1+|t|}.$$

Applying our moment bound in each interval gives

$$\int_{-T}^{T}|\zeta(\frac{1}{2}+it)|^3\frac{\mathrm{d}t}{1+|t|} \ll \sum_{\ell=0}^{\log_2 T}\frac{M_3(T/2^\ell)}{1+T/2^\ell} \ll \sum_{\ell=0}^{\log_2 T}\frac{2^\ell}{T}\frac{T}{2^\ell}(\log T)^{\frac{9}{4}} \ll (\log T)^{\frac{13}{4}}.$$

Finally we look at the integral along the horizontal line segments, these are given by

$$\int_{\frac{1}{2}\pm iT}^{1+\frac{1}{\log x}\pm iT}|F(s)|x^{\mathrm{Re}(s)}\frac{\mathrm{d}s}{|s|} \ll \int_{\frac{1}{2}}^{1+\frac{1}{\log x}}|\zeta(\sigma\pm iT)|^3 x^\sigma\frac{\mathrm{d}\sigma}{T+\sigma}.$$

To bound this we appeal to the Phragmén–Lindelöf principle and the subconvexity bound $|\zeta(\frac{1}{2}+it)| \ll 1+|t|^{\frac{1}{6}}$ due to Weyl. This provides an upper bound of

$$\int_{\frac{1}{2}}^{1+\frac{1}{\log x}}(1+T)^{\frac{1-\sigma}{3}}x^\sigma\frac{\mathrm{d}\sigma}{T+\sigma} \ll (1+T)^{-2/3}\int_{\frac{1}{2}}^{1+\frac{1}{\log x}}(1+T)^{-\sigma/3}x^\sigma\mathrm{d}\sigma$$

which is

$$\ll (1+T)^{-2/3}\left[\frac{x}{T^{1/3}\log(1+T)\log x}+\frac{x^{\frac{1}{2}}}{T^{\frac{1}{2}}\log(1+T)\log x}\right].$$

Our total error is minimised by setting $T=x^3$ which means the principal error term comes from the horizontal strips. □

Applying this we find the count for biquadratic fields to be

$$\frac{1}{6}\sqrt{X}P(\log\sqrt{X})\sum_{C\in\{1,4,8,16\}}\frac{T(C)}{C}+O\left(X^{\frac{1}{4}}\log^{\frac{9}{4}}X\right).$$

It is now a simple check to see that the inner sum is equal to $\frac{23}{8}$, completing the proof.

3.4 Count for fields failing HNP

Similarly to the previous section, we start by making the change of variables

$$m_1=2^{\mu}m_1',\ m_2=\delta_2 2^{\alpha}m_2'\ \text{ and }\ m_3=\delta_3 2^{\beta}m_3', \tag{3.4.1}$$

where $\mu^2(2m_1'm_2'm_3')=1$ and $\mu,\beta,\alpha\in\mathbb{Z}_{\geqslant 0}$ such that $\mu+\alpha+\beta\leqslant 1$. We saw in Section 3.2 that when counting Hasse norm principle failure it is important to keep track of the residue class of $(m_1',m_2',m_3')\bmod 8$ rather than just mod 4 as in Section 3.3. Recall from Section 3.2 that if the congruence classes of m_1,m_2 and $m_3\bmod 4$ are all distinct then K always satisfies the Hasse norm principle. Moreover if exactly two of them are congruent mod 4 then we require that these two are in fact congruent mod 8 to ensure Hasse norm principle failure. We will therefore restrict our attention to $(m_1',\delta_2m_2',\delta_3m_3')$ lying in those admissable residue classes mod 8, denoted $E(\mu,\alpha,\beta)$, for which HNP failure is possible. Specifically

$$E(\mathbf{0})=E_1(\mathbf{0})\cup E_2(\mathbf{0}),$$

where

$$\begin{aligned}E_1(\mathbf{0})&:=\{\boldsymbol{\epsilon}\in((\mathbb{Z}/8\mathbb{Z})^{\times})^3:\epsilon_1\equiv\epsilon_2\equiv\epsilon_3\bmod 4\};\\ E_2(\mathbf{0})&:=\bigcup_{\substack{i,j,k\\ \text{pairwise distinct}}}\{\boldsymbol{\epsilon}\in((\mathbb{Z}/8\mathbb{Z})^{\times})^3:\epsilon_i=\epsilon_j\equiv-\epsilon_k\bmod 4\},\end{aligned}$$

and

$$\begin{aligned}E(1,0,0)&=\{\boldsymbol{\epsilon}\in((\mathbb{Z}/8\mathbb{Z})^{\times})^3:\epsilon_2=\epsilon_3\};\\ E(0,1,0)&=\{\boldsymbol{\epsilon}\in((\mathbb{Z}/8\mathbb{Z})^{\times})^3:\epsilon_1=\epsilon_3\};\\ E(0,0,1)&=\{\boldsymbol{\epsilon}\in((\mathbb{Z}/8\mathbb{Z})^{\times})^3:\epsilon_1=\epsilon_2\}.\end{aligned}$$

Analogously to Section 3.3, we define the constants $C_{\boldsymbol{\delta},\boldsymbol{\epsilon},\mu,\alpha,\beta}$ to account for the different discriminants in each case by setting

$$C_{\boldsymbol{\delta},\boldsymbol{\epsilon},\mu,\alpha,\beta} = \begin{cases} 1 & \text{if } (\epsilon_1, \delta_2\epsilon_2, \delta_3\epsilon_3) \in E_1(\mathbf{0}) \text{ and } \mu = 0 = \alpha = \beta \\ 4 & \text{if } (\epsilon_1, \delta_2\epsilon_2, \delta_3\epsilon_3) \in E_2(\mathbf{0}) \text{ and } \mu = 0 = \alpha = \beta \\ 8 & \text{otherwise.} \end{cases}$$

The counting function for those biquadratic fields of bounded discriminant failing HNP is thus given by

$$\frac{1}{6} \sum_{(\delta_2,\delta_3)\in\{\pm 1\}^2} \sum_{\substack{\mu+\alpha+\beta\in\{0,1\} \\ \mu,\alpha,\beta\in\{0,1\}}} \sum_{(\epsilon_1,\delta_2\epsilon_2,\delta_3\epsilon_3)\in E(\mu,\alpha,\beta)} T(\boldsymbol{\delta},\boldsymbol{\epsilon},\mu,\alpha,\beta), \tag{3.4.2}$$

where $T(\boldsymbol{\delta},\boldsymbol{\epsilon},\mu,\alpha,\beta)$ counts the number of tuples $(m_1', m_2', m_3') \in \mathbb{N}^3$ such that the following all hold:

i) $\mu^2(m_1'm_2'm_3') = 1$,

ii) $(m_1', m_2', m_3') \equiv \boldsymbol{\epsilon} \bmod 8$,

iii) $C_{\boldsymbol{\delta},\boldsymbol{\epsilon},\mu,\alpha,\beta} m_1'm_2'm_3' \leqslant \sqrt{X}$,

iv) (m_1', m_2', m_3') satisfies the local conditions for Hasse norm principle failure in Lemma 3.2.2.

An example of the local conditions necessary for HNP failure is the following

$$p \mid m_1' \implies \left(\frac{m_2m_3}{p}\right) = +1.$$

This condition is strikingly similar to the local expression for the Hilbert symbol as discussed in Chapter 2 and so we can follow the approach of Friedlander–Iwaniec discussed in that chapter. The local condition above is detected using the indicator function given by

$$\prod_{p\mid m_1'} \frac{1}{2}\left(1 + \left(\frac{m_2m_3}{p}\right)\right).$$

Hence we can detect condition (iv) above using

$$\prod_{p\mid m_1'm_2'm_3'} \frac{1}{2}\left(1 + \left(\frac{m_1m_2}{p}\right)\right)\left(1 + \left(\frac{m_1m_3}{p}\right)\right)\left(1 + \left(\frac{m_2m_3}{p}\right)\right).$$

Note that this really is an indicator function because if p divides m_1' then $\left(\frac{m_1m_2}{p}\right) = \left(\frac{m_1m_3}{p}\right) = 0$. Hence only one of the three brackets is distinct from 1 and that bracket must take the value 2 or 0. We next expand this product into a sum

$$\sum_{k_i\mid m_i'} \left(\frac{m_1m_2}{k_3}\right)\left(\frac{m_1m_3}{k_2}\right)\left(\frac{m_2m_3}{k_1}\right).$$

Hence, we have

$$T(\boldsymbol{\delta},\boldsymbol{\epsilon},\mu,\alpha,\beta)=\sum_{\substack{m_1'm_2'm_3'\leqslant M\\ m_i'\equiv\epsilon_i \bmod 8\\ m_i'=k_i\tilde{k}_i}}\frac{\mu^2(m_1'm_2'm_3')}{\tau(m_1'm_2'm_3')}\left(\frac{m_1m_2}{k_3}\right)\left(\frac{m_2m_3}{k_1}\right)\left(\frac{m_3m_1}{k_2}\right),$$

where we have written $M:=\sqrt{X}/C_{\boldsymbol{\delta},\boldsymbol{\epsilon},\mu,\alpha,\beta}$ for the sake of brevity.
Recalling the change of variables (3.3.1), the Legendre symbols are given by

$$\begin{aligned}\left(\frac{m_1m_2}{k_3}\right)&=\left(\frac{2^{\mu+\alpha}\delta_1\delta_2k_1k_2\tilde{k}_1\tilde{k}_2}{k_3}\right),\\ \left(\frac{m_2m_3}{k_1}\right)&=\left(\frac{2^{\alpha+\beta}\delta_2\delta_3k_2k_3\tilde{k}_2\tilde{k}_3}{k_1}\right),\\ \left(\frac{m_3m_1}{k_2}\right)&=\left(\frac{2^{\mu+\beta}\delta_1\delta_3k_1k_3\tilde{k}_1\tilde{k}_3}{k_2}\right).\end{aligned}$$

We may apply quadratic reciprocity to each pair of symbols corresponding to k_i and k_j for $i\neq j$ to conclude

$$T(\boldsymbol{\delta},\boldsymbol{\epsilon},\mu,\alpha,\beta)=\sum_{\substack{m_1'm_2'm_3'\leqslant M\\ m_i'\equiv\epsilon_i \bmod 8\\ m_i'=k_i\tilde{k}_i}}u(\mathbf{k})\frac{\mu^2(m_1'm_2'm_3')}{\tau(m_1'm_2'm_3')}\left(\frac{\tilde{k}_1}{k_2k_3}\right)\left(\frac{\tilde{k}_2}{k_3k_1}\right)\left(\frac{\tilde{k}_3}{k_1k_2}\right),\quad(3.4.3)$$

where $u(\mathbf{k})=(-1)^{\nu(k_1)\nu(k_2)+\nu(k_2)\nu(k_3)+\nu(k_3)\nu(k_1)}\left(\frac{2^\mu}{k_2k_3}\right)\left(\frac{2^\alpha\delta_2}{k_3k_1}\right)\left(\frac{2^\beta\delta_3}{k_1k_2}\right)$. Here $\nu(b)$ is defined to be 0 if $b\equiv 1 \bmod 4$ and 1 otherwise, for any odd integer b.

These sums are now in a form to which the techniques outlined in Chapter 2 may be applied. Specifically we will treat each symbol of the form $\left(\frac{\widetilde{k_i}}{k_j}\right)$ as a character whose modulus is either $k_j,\widetilde{k_i}$ or $4\widetilde{k_i}$ (see e.g. Part 1, Chapter 2, Section 5 of [28]). When the moduli of all of the characters is small then we may apply Lemma 2.2.6 in order to deduce that the main term contribution occurs when all characters are in fact principal. Our first step will be to show that we can always reduce to a situation in which the moduli of all the characters are small.

Let $V=(\log X)^B$ for some large constant parameter B at our disposal. We will show that the main term contribution to $T(\boldsymbol{\delta},\boldsymbol{\epsilon},\mu,\alpha,\beta)$ occurs when the variables lie in one of the following three ranges:

(i) $k_i\leqslant V$ for all i.

(ii) $\tilde{k}_i\leqslant V$ for all i.

(iii) $\tilde{k}_i,\tilde{k}_j,k_i,k_j\leqslant V$ for some choice of $i\neq j$.

Indeed if none of the above scenarios hold then there must exist some $k_i, \widetilde{k_j} > V$ with $i \neq j$. We can exploit cancellation in the sum involving $\left(\frac{\widetilde{k_j}}{k_i}\right)$ using Lemma 2.2.6 to prove that the contribution is negligible. To illustrate this, suppose that $k_2, \tilde{k}_1 > V$. Then we can apply Lemma 2.2.6 to the factor $\left(\frac{\tilde{k}_1}{k_2}\right)$ in (3.4.3). Let

$$f_1(\tilde{k}_1) = \mathbf{1}_{\substack{\tilde{k}_1 \equiv \epsilon_1 k_1 \bmod 8 \\ (\tilde{k}_1, 2k_1\tilde{k}_2k_3\tilde{k}_3)=1}} \frac{\mu^2(\tilde{k}_1)}{\tau(\tilde{k}_1)} \left(\frac{\tilde{k}_1}{k_3}\right)$$
$$f_2(k_2) = \mathbf{1}_{\substack{k_2 \equiv \epsilon_2 \tilde{k}_2 \bmod 8 \\ (k_2, 2k_1\tilde{k}_2k_3\tilde{k}_3)=1}} \frac{\mu^2(k_2)}{\tau(k_2)} \left(\frac{\tilde{k}_3}{k_2}\right).$$

Then the k_2, $\tilde{k}_1$ sum is given by

$$\sum_{\substack{k_2, \tilde{k}_1 \geqslant V \\ k_2\tilde{k}_1 \leqslant \frac{\sqrt{X}}{k_1\tilde{k}_3k_3\tilde{k}_2} \\ (\tilde{k}_1, k_2)=1}} f_1(\tilde{k}_1) f_2(k_2) u(\mathbf{k}) \left(\frac{\tilde{k}_1}{k_2}\right) \ll \sqrt{X} V^{-\frac{1}{6}} (\log X)^{\frac{7}{6}} \frac{1}{\tilde{k}_2 k_1 k_3 \tilde{k}_3}.$$

Repeatedly applying the estimate $\sum_{d \leqslant x} \frac{1}{d\tau(d)} \ll \sqrt{\log x}$ gives a bound for the sum over the remaining variables of

$$\ll \sqrt{X} V^{-\frac{1}{6}} (\log X)^{\frac{19}{6}}. \tag{3.4.4}$$

Of course, the bound (3.4.4) applies for all ranges $k_i, \tilde{k}_j > V$ where $i \neq j$.

This allows us to restrict the variables to lie in one of the ranges (i), (ii) or (iii). However in fact range (iii) also provides a negligible contribution. Indeed suppose $k_1, \widetilde{k_1}, k_2, \widetilde{k_2} \leqslant V$ then the contribution to $T(\boldsymbol{\delta}, \boldsymbol{\epsilon}, \mu, \alpha, \beta)$ can be bounded above by

$$\sum_{\substack{m_1'm_2'm_3' \leqslant M \\ m_1', m_2' \leqslant V^2}} \frac{\mu^2(m_1'm_2'm_3')}{\tau(m_1'm_2'm_3')} \ll \frac{\sqrt{X}}{\sqrt{\log X}} \sum_{m_1', m_2' \leqslant V^2} \frac{\mu^2(m_1'm_2')}{\tau(m_1'm_2')m_1'm_2'}$$
$$\ll \frac{\sqrt{X} \log V}{\sqrt{\log X}}.$$

For the remainder of this section then we may assume that we are in one of the ranges (i) or (ii) described above. In the first case where $k_i \leqslant V$ for all i, define

$$S(k_1, k_2, k_3) = \sum_{\substack{\widetilde{k_1}\widetilde{k_2}\widetilde{k_3} \leqslant M/k_1k_2k_3 \\ k_i \equiv \epsilon_i \widetilde{k_i} \bmod 8 \\ \gcd(\widetilde{k_1}\widetilde{k_2}\widetilde{k_3}, 2k_1k_2k_3)=1}} \frac{\mu^2(\widetilde{k_1}\widetilde{k_2}\widetilde{k_3})}{\tau(\widetilde{k_1}\widetilde{k_2}\widetilde{k_3})} \left(\frac{\tilde{k}_1}{k_2k_3}\right)\left(\frac{\tilde{k}_2}{k_3k_1}\right)\left(\frac{\tilde{k}_3}{k_1k_2}\right).$$

To compute the contribution of this range to the total sum $T(\boldsymbol{\delta}, \boldsymbol{\epsilon}, \mu, \alpha, \beta)$ we must evaluate

$$\sum_{\substack{k_1k_2k_3 \leqslant M \\ 2 \nmid k_1k_2k_3}} \frac{\mu^2(k_1k_2k_3)}{\tau(k_1k_2k_3)} u(\mathbf{k}) S(k_1, k_2, k_3).$$

Similarly the contribution from the second range where we have $\widetilde{k_i} \leqslant V$ is given by

$$\sum_{\widetilde{k_1}\widetilde{k_2}\widetilde{k_3} \leqslant M} \frac{\mu^2(k_1k_2k_3)}{\tau(k_1k_2k_3)} u(\epsilon_1\widetilde{k_1}, \epsilon_2\widetilde{k_2}, \epsilon_3\widetilde{k_3}) \widetilde{S}(\widetilde{k_1}, \widetilde{k_2}, \widetilde{k_3}),$$

where

$$\widetilde{S}(\widetilde{k_1}, \widetilde{k_2}, \widetilde{k_3}) = \sum_{\substack{k_1k_2k_3 \leqslant M/\widetilde{k_1}\widetilde{k_2}\widetilde{k_3} \\ \widetilde{k_i} \equiv \epsilon_i k_i \bmod 8 \\ \gcd(k_1k_2k_3, 2\widetilde{k_1}\widetilde{k_2}\widetilde{k_3})=1}} \frac{\mu^2(k_1k_2k_3)}{\tau(k_1k_2k_3)} \left(\frac{\widetilde{k_1}\widetilde{k_2}}{k_3}\right)\left(\frac{\widetilde{k_1}\widetilde{k_3}}{k_2}\right)\left(\frac{\widetilde{k_2}\widetilde{k_3}}{k_1}\right).$$

Note here that we have implicitly used the fact that the reciprocity factor $u(\mathbf{k})$ is completely determined by the residue class of $\mathbf{k}$ mod 8. Indeed $\nu(x)$ is only dependent on the residue class of x mod 4 and $\left(\frac{2}{x}\right)$ is determined by the residue class of x mod 8. Thus $u(\mathbf{k}) = u(\epsilon_1\widetilde{k_1}, \epsilon_2\widetilde{k_2}, \epsilon_3\widetilde{k_3})$.

We are now ready to put Lemma 2.2.4 to use and compute the values of S and $\widetilde{S}$.

Lemma 3.4.1. *If $k_i \leqslant V$ for all i then for any $A > 0$ we have*

$$S(k_1, k_2, k_3) = \frac{M}{56\sqrt{\pi}} \delta_{\mathbf{k}} \prod_p \left(1 + \frac{3}{2p}\right)\left(1 - \frac{1}{p}\right)^{\frac{3}{2}} \left\{\sqrt{\log M} + O(1)\right\} + O_A\left(\frac{M}{(\log M)^A}\right),$$

where $\delta_{\mathbf{k}} = 1$ if each $k_i = 1$ and 0 otherwise. By symmetry, the same formula holds for $\widetilde{S}(\widetilde{k_1}, \widetilde{k_2}, \widetilde{k_3})$ in the range $\widetilde{k_i} \leqslant V$ for all i.

Proof. We will focus just on the case when $k_i \leqslant V$. We will show that if any of the characters $\left(\frac{\cdot}{k_i}\right)$ is non-principal, we get a negligible contribution. Indeed suppose that $\left(\frac{\cdot}{k_1}\right)$ is non-principal. Applying the error term in Lemma 2.2.4 we have, for any $A > 0$, the bound

$$S(k_1, k_2, k_3) \ll_A \sum_{\widetilde{k_2}\widetilde{k_3} \leqslant M} \frac{\mu^2(\widetilde{k_2}\widetilde{k_3})}{\tau(\widetilde{k_2}\widetilde{k_3})} \tau(\widetilde{k_2}\widetilde{k_3}k_1k_2k_3) \frac{M}{\widetilde{k_2}\widetilde{k_3}(\log M)^A}.$$

Performing the $\widetilde{k_2}, \widetilde{k_3}$ sum trivially gives

$$S(k_1, k_2, k_3) \ll_A \frac{M\tau(k_1k_2k_3)}{(\log M)^A}.$$

The total contribution then from this range is given by

$$\ll_A \frac{MV^3}{(\log M)^A}.$$

Therefore the main term contribution occurs when each of the characters $\left(\frac{\cdot}{k_i}\right)$ are principal. This can only occur when $k_1 = k_2 = k_3 = 1$. In this case,

$$S(1,1,1) = \sum_{\substack{\widetilde{k_1}\widetilde{k_2}\widetilde{k_3} \leqslant M \\ k_i \equiv \epsilon_i \bmod 8}} \frac{\mu^2(\widetilde{k_1}\widetilde{k_2}\widetilde{k_3})}{\tau(\widetilde{k_1}\widetilde{k_2}\widetilde{k_3})}.$$

The analysis of this sum follows very closely the proof in Section 3.3. Firstly we remove the congruence condition by summing over characters mod 8. Thus

$$S(1,1,1) = \frac{1}{4^3} \sum_{\chi_i \bmod 8} \prod_{i=1}^{3} \overline{\chi_i(\epsilon_i)} \sum_{\widetilde{k_1}\widetilde{k_2}\widetilde{k_3} \leqslant M} \frac{\mu^2(\widetilde{k_1}\widetilde{k_2}\widetilde{k_3})}{\tau(\widetilde{k_1}\widetilde{k_2}\widetilde{k_3})} \chi_1(\widetilde{k_1})\chi_2(\widetilde{k_2})\chi_3(\widetilde{k_3}).$$

Again we look at the associated Dirichlet series given by

$$\sum_{n \text{ odd}} \frac{\mu^2(n)}{\tau(n)n^s} \sum_{n_1 n_2 n_3 = n} \chi_1(n_1)\chi_2(n_2)\chi_3(n_3).$$

This series is equal to $\zeta(s)^k \widetilde{G}(s)$ where $\widetilde{G}(s)$ is absolutely convergent in the region $\mathrm{Re}(s) > \frac{1}{2}$ and $k = \frac{3 - \#\{i : \chi_i \text{ non-principal}\}}{2}$. If any of the χ_i are non-principal then its Dirichlet L-functions is entire and trivially satisfies the same subconvexity and moment bounds as $\zeta(s)$. Therefore an application of Perron's formula as in the proof of Lemma 3.3.1 yields a contribution of $O(M)$. The leading contribution occurs in the case that all χ_i are principal. Here we are looking at the sum

$$\Sigma(M) = \sum_{\widetilde{k_1}\widetilde{k_2}\widetilde{k_3} \leqslant M} \frac{\mu^2(\widetilde{k_1}\widetilde{k_2}\widetilde{k_3})}{\tau(\widetilde{k_1}\widetilde{k_2}\widetilde{k_3})} = \sum_{\substack{n \leqslant M \\ n \text{ odd}}} \mu^2(n) \left(\frac{3}{2}\right)^{\omega(n)}.$$

In analogy with Lemma 3.3.1, we have

$$\Sigma(M) := \frac{1}{\Gamma(3/2)} \left(1 + \frac{3}{2 \times 2}\right)^{-1} M\sqrt{\log M} \prod_p \left(1 + \frac{3}{2p}\right)\left(1 - \frac{1}{p}\right)^{\frac{3}{2}} + O(M).$$

□

All that is left to complete our analysis of (3.4.2) is to perform the sums

$$\sum_{(\delta_2,\delta_3) \in \{\pm 1\}^2} \sum_{\substack{\mu+\alpha+\beta \in \{0,1\} \\ \mu,\alpha,\beta \in \{0,1\}}} \sum_{(\epsilon_1, \delta_2\epsilon_2, \delta_3\epsilon_3) \in E(\mu,\alpha,\beta)} \frac{1}{C_{\boldsymbol{\delta},\boldsymbol{\epsilon},\mu,\alpha,\beta}},$$

and

$$\sum_{(\delta_2,\delta_3)\in\{\pm1\}^2}\sum_{\substack{\mu+\alpha+\beta\in\{0,1\}\\ \mu,\alpha,\beta\in\{0,1\}}}\sum_{(\epsilon_1,\delta_2\epsilon_2,\delta_3\epsilon_3)\in E(\mu,\alpha,\beta)}\frac{u(\epsilon_1,\epsilon_2,\epsilon_3)}{C_{\boldsymbol{\delta},\boldsymbol{\epsilon},\mu,\alpha,\beta}}.$$

Lemma 3.4.2. *We have*

$$\sum_{(\delta_2,\delta_3)\in\{\pm1\}^2}\sum_{\substack{\mu+\alpha+\beta\in\{0,1\}\\ \mu,\alpha,\beta\in\{0,1\}}}\sum_{(\epsilon_1,\delta_2\epsilon_2,\delta_3\epsilon_3)\in E(\mu,\alpha,\beta)}\frac{u(\epsilon_1,\epsilon_2,\epsilon_3)}{C_{\boldsymbol{\delta},\boldsymbol{\epsilon},\mu,\alpha,\beta}}=0.$$

Proof. We start by observing that

$$\left(\frac{2^\mu}{\epsilon_2\epsilon_3}\right)\left(\frac{2^\alpha}{\epsilon_1\epsilon_3}\right)\left(\frac{2^\beta}{\epsilon_1\epsilon_2}\right)=1.$$

Indeed this is clearly true when $\mu+\alpha+\beta=0$. If $\mu=1$ then by the definition of $E(1,0,0)$ we must have $\delta_2\epsilon_2=\delta_3\epsilon_3$ therefore

$$\left(\frac{2}{\epsilon_2\epsilon_3}\right)=1.$$

The other cases follow similarly.
Therefore

$$\begin{aligned}u(\boldsymbol{\epsilon})&=(-1)^{\nu(\epsilon_1)\nu(\epsilon_2)+\nu(\epsilon_3)\nu(\epsilon_1)+\nu(\epsilon_2)\nu(\epsilon_3)}\left(\frac{\delta_2}{\epsilon_1\epsilon_3}\right)\left(\frac{\delta_3}{\epsilon_1\epsilon_2}\right)\\&=(-1)^{\nu(\epsilon_1)\nu(\epsilon_2)+\nu(\epsilon_3)\nu(\epsilon_1)+\nu(\epsilon_2)\nu(\epsilon_3)+\nu(\delta_2)\nu(\epsilon_1\epsilon_3)+\nu(\delta_3)\nu(\epsilon_1\epsilon_2)}\end{aligned}$$

Now we just run through all possible values of $\delta_2,\delta_3,\boldsymbol{\epsilon},\mu,\alpha$ and β and see what comes out. For simplicity we write $\boldsymbol{\epsilon}':=(\epsilon_1,\delta_2\epsilon_2,\delta_3\epsilon_3)$.

First suppose that $(\delta_2,\delta_3)=(+1,+1)$ then

$$u(\boldsymbol{\epsilon})=(-1)^{\nu(\epsilon_1)\nu(\epsilon_2)+\nu(\epsilon_3)\nu(\epsilon_1)+\nu(\epsilon_2)\nu(\epsilon_3)}.$$

By Lemma 3.2.2, we know that at least two components of $\boldsymbol{\epsilon}'$ must be equal therefore for some $\mu\in(\mathbb{Z}/8\mathbb{Z})^\times$ we have

$$u(\boldsymbol{\epsilon})=(-1)^\mu$$

Hence the sum over these is 0.

Now suppose $(\delta_2,\delta_3)=(+1,-1)$ so that

$$u(\boldsymbol{\epsilon})=(-1)^{\nu(\epsilon_1)\nu(\epsilon_2)+\nu(\epsilon_3)\nu(\epsilon_1)+\nu(\epsilon_2)\nu(\epsilon_3)+\nu(\epsilon_1\epsilon_2)}.$$

If $\boldsymbol{\epsilon}'\in E_1(\mathbf{0})$ then $\epsilon_1\equiv\epsilon_2\equiv-\epsilon_3 \bmod 4$ so

$$u(\boldsymbol{\epsilon})=(-1)^{\nu(\epsilon_1)}.$$

So again this sums to 0. Next suppose $\boldsymbol{\epsilon}'\in E_2(\mathbf{0})$ then one of the following cases occurs

(i) $\epsilon_1 = \epsilon_2 \equiv -\epsilon_3 \bmod 4$ then $u(\boldsymbol{\epsilon}) = (-1)^{\nu(\epsilon_1)+1}$.

(ii) $\epsilon_2 = \epsilon_3 \equiv -\epsilon_1 \bmod 4$ then $u(\boldsymbol{\epsilon}) = (-1)^{\nu(\epsilon_1)+1}$.

(iii) $\epsilon_3 = \epsilon_1 \equiv -\epsilon_2 \bmod 4$ then $u(\boldsymbol{\epsilon}) = (-1)^{\nu(\epsilon_1)}$.

Each of these cases sums to 0.

If $\boldsymbol{\epsilon}' \in E(1,0,0)$ then $\epsilon_2 = -\epsilon_3$ so

$$u(\boldsymbol{\epsilon}) = (-1)^{\nu(\epsilon_1)+\nu(\epsilon_1\epsilon_2)}.$$

If $\boldsymbol{\epsilon}' \in E(0,1,0)$ then $\epsilon_1 = -\epsilon_3$ so

$$u(\boldsymbol{\epsilon}') = (-1)^{\nu(\epsilon_2)+\nu(\epsilon_1\epsilon_2)}.$$

If $\boldsymbol{\epsilon}' \in E(0,0,1)$ then $\epsilon_1 = \epsilon_2$ so

$$u(\boldsymbol{\epsilon}) = (-1)^{\nu(\epsilon_1)}.$$

Again, all of these sum to 0. The case where $(\delta_2, \delta_3) = (-1, +1)$ is similar.

Finally suppose $(\delta_2, \delta_3) = (-1, -1)$, in which case

$$u(\boldsymbol{\epsilon}) = (-1)^{\nu(\epsilon_1)\nu(\epsilon_2)+\nu(\epsilon_3)\nu(\epsilon_1)+\nu(\epsilon_2)\nu(\epsilon_3)+\nu(\epsilon_1\epsilon_2)+\nu(\epsilon_1\epsilon_3)}.$$

Then for $\boldsymbol{\epsilon}' \in E_1(\mathbf{0})$ we have

$$u(\boldsymbol{\epsilon}) = (-1)^{\nu(-\epsilon_1)}.$$

For $\boldsymbol{\epsilon}' \in E(1,0,0)$ we must have $\epsilon_2 = \epsilon_3$ so

$$u(\boldsymbol{\epsilon}) = (-1)^{\nu(\epsilon_2)}.$$

For $\boldsymbol{\epsilon}' \in E(0,1,0)$ we must have $\epsilon_1 = -\epsilon_3$ so

$$u(\boldsymbol{\epsilon}) = (-1)^{\nu(\epsilon_2)+\nu(\epsilon_1\epsilon_2)}.$$

For $\boldsymbol{\epsilon}' \in E(0,0,1)$ we must have $\epsilon_1 = -\epsilon_2$ so

$$u(\boldsymbol{\epsilon}) = (-1)^{\nu(\epsilon_3)+\nu(\epsilon_1\epsilon_3)}.$$

In all of these cases the sum is 0. □

It is merely a combinatorial exercise to see that

$$\sum_{(\delta_2,\delta_3)\in\{\pm1\}^2} \sum_{\substack{\mu+\alpha+\beta\in\{0,1\}\\ \mu,\alpha,\beta\in\{0,1\}}} \sum_{(\epsilon_1,\delta_2\epsilon_2,\delta_3\epsilon_3)\in E(\mu,\alpha,\beta)} \frac{1}{C_{\boldsymbol{\delta},\boldsymbol{\epsilon},\mu,\alpha,\beta}} = 112. \tag{3.4.5}$$

Combining (3.4.2), Lemma 3.4.1 and (3.4.5) yields the claimed main term.

3.5 Ordering by conductor

We include here the proof of Theorem 3.1.4 which is a simple corollary of the previous two sections. By the Kronecker–Weber theorem every abelian number field lies inside a cyclotomic field $\mathbb{Q}(\zeta_N)$. The conductor C_K of an abelian number field $K/\mathbb{Q}$ is defined to be the least N such that $K \subset \mathbb{Q}(\zeta_N)$. The conductor-discriminant formula gives an explicit description of the discriminant and conductor of a number field in terms of the Artin conductors of irreducible characters on the Galois group. The following two expressions are consequences of this description.

Lemma 3.5.1. *Suppose that $K/\mathbb{Q}$ is a biquadratic number field with k_1, k_2, k_3 its 3 distinct quadratic subfields. Then*

$$C_K = lcm\{C_{k_1}, C_{k_2}, C_{k_3}\}$$

and

$$C_{k_i} = |\Delta_{k_i}|.$$

Therefore the conductor of the biquadratic field $\mathbb{Q}(\sqrt{m_1m_2}, \sqrt{m_1m_3})$ is $2^\delta m_1m_2m_3$ where $\delta = 0$ if $m_i \equiv 1 \bmod 4$ for all i, and 1 otherwise.

To count the number of fields of bounded conductor, in analogy with (3.3.2), we must compute

$$\frac{1}{6}\sum_{c\in\{1,4,8\}} \widetilde{T}(c) \sum_{\substack{m_1'm_2'm_3' \leqslant X/c \\ 2\nmid m_1'm_2'm_3'}} \mu^2(m_1'm_2'm_3'),$$

where we have defined the constants c analagously to in the previous sections. i.e.

$$c = \begin{cases} 1 \text{ if } \mu = \alpha = \beta = 0 \text{ and } m_1 \equiv m_2 \equiv m_3 \equiv 1 \bmod 4, \\ 4 \text{ if } \mu = \alpha = \beta = 0 \text{ and } \exists i \text{ such that } m_i \not\equiv 1 \bmod 4 \\ 8 \text{ otherwise.} \end{cases}$$

By Lemma 3.3.1, the inner sum is

$$\frac{X}{c}P(\log X) + O(X^{\frac{1}{2}}\log^{\frac{13}{4}} X).$$

Computing the sum $\sum_{c\in\{1,4,8\}} \frac{\widetilde{T}(c)}{c} = \frac{13}{4}$ gives the field count claimed in Theorem 3.1.4. We note that this asymptotic formula is in agreement with the one proven by Mäki [83, Theorem 3]. Indeed, Mäki's formula for the constant in the biquadratic case is

$$\frac{1}{6}\left(\frac{3}{4} + 2^{2-2} + 2^{2-3}3\right)c_0,$$

where

$$c_0 = \frac{1}{2!}\prod_{p\neq 2}\left(1+\frac{3}{p}\right)\left(1-\frac{1}{p}\right)^3(1-2^{-1})^3.$$

The error term in [83] is $O(X^{\frac{2}{3}})$, which is the equivalent of the error term in [12], and our proof produces an improvement in this aspect.

We can set up the count for Hasse norm principle failures in the exact same way as in (3.4.2) except that now we wish to count those triples (m_1', m_2', m_3') such that $m_1'm_2'm_3' \leqslant cX$. Let us denote this quantity by $\widetilde{T}(\boldsymbol{\delta}, \boldsymbol{\epsilon}, \mu, \alpha, \beta)$. By Lemma 3.4.2 we have that

$$\widetilde{T}(\boldsymbol{\delta}, \boldsymbol{\epsilon}, \mu, \alpha, \beta) = \frac{X\sqrt{\log X}}{c56\sqrt{\pi}}\left(1 + u(\boldsymbol{\epsilon})\right)\prod_{p}\left(1 + \frac{3}{2p}\right)\left(1 - \frac{1}{p}\right)^{\frac{3}{2}} + O(X).$$

We have again the same cancellation in the sum involving $u(\boldsymbol{\epsilon})$ as in Lemma 3.4.2. Finally we note that by the definition of c and of $C_{\boldsymbol{\delta},\boldsymbol{\epsilon},\mu,\alpha,\beta}$ the remaining sum is precisely (3.4.5), from which the result follows.

Chapter 4

Hasse Principle Failures in a Family of Châtelet Surfaces

The material of this chapter is based on [98].

4.1 Introduction

As discussed in Examples 1.2.19 and 1.3.7 in the introductory chapter, a Châtelet surface over a number field k is a smooth proper model of the affine surface defined by the equation

$$Y^2 - eZ^2 = f(T),$$

where $e \in K^*$ is non-square and $f \in k[T]$ is a separable polynomial of degree 3 or 4. These varieties have featured heavily in the study of rational points since they are rational surfaces which are neither the projective plane nor a quadric and which naturally admit a conic bundle structure. Moreover the left hand side of the defining equation is a quadratic norm. Since the representation function of a number by such a norm is a well studied arithmetic function, the application of analytic techniques to counting rational points on these surfaces has lead to a great deal of progress on Manin's programme.

The work of Colliot-Thélène, Sansuc and Swinnerton-Dyer [35], in their own words, "solves all sensible Diophantine problems regarding the rational points of Châtelet surfaces". That is to say they were able to show that the Hasse principle holds unless f is the product of two irreducible quadratics and that the Brauer–Manin obstruction is the only obstruction to the existence of rational points and to weak approximation. We restrict our attention to the case $k = \mathbb{Q}$ and the Châtelet surfaces $X_{a,b,c,d}$, defined by the equation

$$Y^2 + Z^2 = (aT^2 + b)(cT^2 + d) \neq 0, \tag{4.1.1}$$

for $(a, b, c, d) \in \mathbb{Q}^4$ such that $abcd \neq 0$ and $ad - bc \neq 0$. There exist several explicit counterexamples to the Hasse principle for surfaces of this type in the

literature. Colliot-Thélène, Coray and Sansuc [31, Proposition C] provide an infinite family of counterexamples over $\mathbb{Q}$ given by $X_{1,1-k,-1,k}$ for any positive integer $k \equiv 3 \bmod 4$. This generalises the earlier example (Chapter 1, Example 1.2.19) of Iskovskikh [66] where $k = 3$. An investigation of the surfaces defined by (4.1.1) was undertaken by la Bretèche and Browning [13], in which they vary the 4-tuple of coefficients (a, b, c, d) and develop asymptotics for what proportion produce counterexamples to the Hasse principle. The main results of [13] show that roughly 83.3% of Châtelet surfaces have local solutions everywhere but 0% are counterexamples to the Hasse principle.

Remark. This family admits an obvious fibration (1.3.3) over an open subset of $\mathbb{P}^3$ by projecting onto the coefficients. Each fibre over a codimension one point is split and so the Loughran–Smeets exponent $\Delta(\pi)$ is 0, making the la Bretèche–Browning local result a special case of the theorem of Loughran and Smeets.

The purpose of this chapter is to discover a large family of surfaces among which a positive proportion fail the Hasse principle. Our principle tool will be a direct study of the Brauer–Manin obstruction for such surfaces, which we know to be the only one. Further, in the case of surfaces $X = X_{a,b,c,d}$ of the form (4.1.1), we have $\operatorname{Br} X / \operatorname{Br} \mathbb{Q} \cong \mathbb{Z}/2\mathbb{Z}$ and that the quotient is generated by the quaternion algebra $(-1, aT^2 + b)$. Hence the Brauer–Manin obstruction can be made very explicit. Given $(a, b, c, d) \in \mathbb{Q}^4$, we have

$$X(\mathbb{A}_{\mathbb{Q}})^{\mathrm{Br}} = \left\{ (x_\nu, y_\nu, t_\nu)_{\nu \in \Omega} \in X(\mathbb{A}_{\mathbb{Q}}) : \prod_{\nu \in \Omega} \left(\frac{-1, at_\nu^2 + b}{\mathbb{Q}_\nu} \right) = +1 \right\},$$

where Ω denotes the set of places of $\mathbb{Q}$ and we recall that $\left(\frac{\cdot,\cdot}{\mathbb{Q}_\nu}\right)$ denotes the Hilbert symbol associated to the field $\mathbb{Q}_\nu$. Since the Brauer–Manin obstruction is the only one $X(\mathbb{Q}) = \varnothing \iff X(\mathbb{A}_{\mathbb{Q}})^{\mathrm{Br}} = \varnothing$, the surface $X_{a,b,c,d}$ is a counterexample to the Hasse principle exactly when the above product of Hilbert symbols is -1 for any point $(x_\nu, y_\nu, t_\nu)_{\nu \in \Omega} \in X(\mathbb{A}_{\mathbb{Q}})$. We will see in Section 4.3 that a way to do this is to impose the condition $|ad - bc| = 1$ which will force the Hilbert symbol $\left(\frac{-1, at_p^2 + b}{\mathbb{Q}_p}\right)$ to be constant at all odd primes p. Note that this condition is satisfied by surfaces in the Colliot-Thélène–Coray–Sansuc family mentioned above.

Remark. One could instead impose the condition $|ad - bc| = p$ for a prime p and the local Hilbert symbol would be constant for all but at most 2 places which would again lead to a positive proportion of Hasse principle failures. Indeed further one could impose the condition that $ad - bc$ had at most some fixed number of prime factors. We do not pursue this extension further as it would follow the same process and serve only to make the exposition messier.

Before counting, we make some initial reduction steps in order to uniquely parametrise (upto isomorphism) the surfaces in our family. We will identify (a, b, c, d) with

$(-a,-b,-c,-d)$ as this does not affect the surface in any way. Clearly switching the order of the two brackets on the right hand side of equation (4.1.1) will not change the equation either therefore we see that $X_{a,b,c,d}$ is invariant under the map

$$\varrho_1 : (a,b,c,d) \mapsto (c,d,a,b).$$

Moreover, since the surface is a projective variety, it makes no difference if we divide by T in (4.1.1) and relabel Y and Z accordingly. Therefore $X_{a,b,c,d}$ is isomorphic to the surface obtained by

$$\varrho_2 : (a,b,c,d) \mapsto (b,a,d,c).$$

For the same reason, $X_{a,b,c,d}$ is equivalent to $X_{\lambda a,\lambda b,\lambda c,\lambda d}$ for any $\lambda \in \mathbb{Q}^*$. Each Châtelet surface of the form (4.1.1) can be written as $X_{a,b,c,d}$ where $(a,b,c,d) \in \mathbb{Z}^4_{\text{prim}}$. For such a tuple (a,b,c,d) the expression $|ad-bc| = 1$ is invariant under ϱ_1 and ϱ_2. We introduce the following set which will act as our parameter space for Châtelet surfaces

$$\mathcal{M} := \{(a,b,c,d) \in \mathbb{Z}^4/\{\pm 1\} : |ad-bc| = 1,\, abcd \neq 0\}/\sim,$$

where two tuples are equivalent under $\sim$ if they are in the same orbit of ϱ_1 and ϱ_2. We will be interested in the surfaces which have local solutions everywhere and those which have local solutions everywhere but no global solutions. These sets will be denoted $\mathcal{M}_{\text{loc}}$ and $\mathcal{M}_{\text{Br}}$, respectively. On choosing the sup-norm for points $u = (a,b,c,d)$ on $\mathcal{M}$, i.e. $|u| = \max\{|a|,|b|,|c|,|d|\}$, we can define the counting functions, analogously to Section 1.3, by

$$N_{\text{loc/Br}}(B) = \#\{u \in \mathcal{M}_{\text{loc/Br}} : |u| \leqslant B\}. \tag{4.1.2}$$

The object of this chapter is to establish the following asymptotic formulae for these counting functions.

Theorem 4.1.1. *There exist $\theta, \theta' > 0$ such that*

$$N_{loc}(B) = \frac{279}{16\pi^2} B^2 + O\left(B^{2-\theta}\right) \tag{4.1.3}$$

and

$$N_{Br}(P) = \frac{33}{8\pi^2} B^2 + O\left(B^{2-\theta'}\right). \tag{4.1.4}$$

Remark. The constants above should be considered as products of local densities (c.f. Lemma 4.4.2). In the case of N_{loc} this is in line with the result of Loughran–Smeets. For N_{Br} there is no reason to expect this to be true.

Denoting by $N(P)$ the total number of surfaces in this family, it is established in Theorem 4.4.1 that

$$N(B) \sim \frac{32}{\pi^2} B^2.$$

Theorem 4.1.1 therefore shows that a positive proportion of surfaces of this type fail the Hasse principle.

4.2 Set-up

We will choose as a set of representatives for $\mathcal{M}$ the following set,

$$S_{\text{tot}} = \{(a,b,c,d) \in \mathbb{Z}^4 : a > 0, |ad-bc| = 1, abcd \neq 0\}. \tag{4.2.1}$$

Observe that each tuple in $\mathcal{M}$ is counted by 4 distinct elements of S_{tot} due to the invariance under ϱ_1 and ϱ_2. This means there is a 1:4 correspondence between $\mathcal{M}$ and S_{tot}. It will be important later on to keep track of the 2-adic valuation of the coefficients. In order to simplify this, we will restrict to representatives for which a is odd. We may do this by breaking up S_{tot} as

$$S_{\text{tot}}^{(1)} \sqcup S_{\text{tot}}^{(2)}$$

where $S_{\text{tot}}^{(1)} := \{(a,b,c,d) \in S_{\text{tot}} : 2 \nmid a\}$ and $S_{\text{tot}}^{(2)} := \{(a,b,c,d) \in S_{\text{tot}} : 2 \mid a\}$. Due to the expression $|ad-bc| = 1$ we know that when 2 divides a, it cannot divide b. Therefore by the invariance of Châtelet surfaces under the action of ϱ_2 we may replace the previous expression for $S_{\text{tot}}^{(2)}$ by $S_{\text{tot}}^{(2)} = \{(a,b,c,d) \in S_{\text{tot}} : 2 \nmid a, 2 \mid b\}$. For our purposes, we are interested in those elements of S_{tot} that define surfaces $X = X_{a,b,c,d}$ with points everywhere locally, especially those without points globally. To this end, we observe that the counting functions defined in (4.1.2) may be written as

$$N_{\text{loc}}(B) = \frac{1}{4}\#\{t = (a,b,c,d) \in S_{\text{tot}} : |t| \leqslant B, X(\mathbb{Q}_\nu) \neq \varnothing \forall \nu \in \Omega\}$$
$$N_{\text{Br}}(B) = \frac{1}{4}\#\{t = (a,b,c,d) \in S_{\text{tot}} : |t| \leqslant B, X(\mathbb{Q}_\nu) \neq \varnothing \forall \nu \in \Omega, X(\mathbb{Q}) = \varnothing\}.$$

4.3 Brauer group considerations

In this section we aim to work with elements of the Brauer group of the surfaces in our family to show that the Brauer–Manin obstruction is completely controlled by what happens at the prime 2. Before doing so we note that it is possible to obtain Theorem4.1.1 by instead following the work of la Bretèche and Browning more closely. In [13], they develop explicit conditions on the parameters (a,b,c,d) that guarantee local solubility of the surface $X_{a,b,c,d}$. By investigating a related torsor they are then able to construct similar conditions controlling global solubility. Summing over the (a,b,c,d) that satisfy the local conditions but not the global conditions allows them to develop their asymptotics. One could specialise their results to the case when $|ad-bc| = 1$ however we choose instead to work directly with the Brauer group, for the sake of brevity, to tie in with other discussions of the Brauer–Manin obstruction in this thesis and so that we need not make appeal to the high powered machinery of descent on torsors.

Lemma 4.3.1. *Let $(a,b,c,d) \in S_{tot}$. If p is an odd prime then we have*

$$\left(\frac{-1, at_p^2 + b}{\mathbb{Q}_p}\right) = +1,$$

for any $(x_p, y_p, t_p) \in X_{a,b,c,d}(\mathbb{Q}_p)$.

Proof. This is immediate for $p \equiv 1 \bmod 4$. If $p \equiv 3 \bmod 4$ then

$$\left(\frac{-1, at_p^2 + b}{\mathbb{Q}_p}\right) = (-1)^{v_p(at_p^2+b)}.$$

Note that since

$$x_p^2 + y_p^2 = (at_p^2 + b)(ct_p^2 + d),$$

we have

$$(-1)^{v_p(at_p^2+b)} = (-1)^{v_p(ct_p^2+d)}.$$

Now observe that

$$a(ct_p^2 + d) - c(at_p^2 + b) = ad - bc = \pm 1.$$

Hence either $v_p(at_p^2 + b) = 0$ or $v_p(ct_p^2 + d) = 0$. □

We know turn to investigating the real place. Recall that $a > 0$ for all $(a, b, c, d) \in S_{\text{tot}}$. There are four possible combinations of signs that the remaining coefficients b, c and d can take while still maintaining the relationship $ad - bc = \pm 1$. In particular,

$$\sigma(b, c, d) \in \{(+, +, +), (+, -, -), (-, +, -), (-, -, +)\},$$

where σ is the 3 dimensional version of the usual sign function

$$\sigma(x) = \begin{cases} + \text{ if } x > 0 \\ - \text{ if } x < 0. \end{cases}$$

The third case is the subject of Lemma 4.5 in [13] wherein it is shown that surfaces defined by coefficients with this signature always satisfy the Hasse principle. In all other cases, the Hilbert symbol is constant.

Lemma 4.3.2. *Let* $(a, b, c, d) \in S_{tot}$ *such that* $\sigma(b, c, d) \neq (-, +, -)$. *Then the Hilbert symbol* $\left(\frac{-1, at_\infty^2 + b}{\mathbb{R}}\right)$ *is constant as a function of* $(x_\infty, y_\infty, t_\infty) \in X_{a,b,c,d}(\mathbb{R})$.

Proof. The definition of the real Hilbert symbol tells us that

$$\left(\frac{-1, at_\infty^2 + b}{\mathbb{R}}\right) = +1 \iff at_\infty^2 + b > 0.$$

Since $a > 0$, if $b > 0$ we immediately have that $\left(\frac{-1, at_\infty^2 + b}{\mathbb{R}}\right) = +1$. It remains to investigate the last case $\sigma(b, c, d) = (-, -, +)$. In this case, we claim that

$$\left(\frac{-1, at_\infty^2 + b}{\mathbb{R}}\right) = ad - bc.$$

Let $(x_\infty, y_\infty, t_\infty) \in X_{a,b,c,d}(\mathbb{R})$ and suppose that $\left(\frac{-1, at_\infty^2+b}{\mathbb{R}}\right) = -1$. Since

$$x_\infty^2 + y_\infty^2 = (at_\infty^2 - |b|)(d - |c|t_\infty^2),$$

we have $at_\infty^2 - |b| < 0$ and $d - |c|t_\infty^2 < 0$. Hence

$$\frac{d}{|c|} < t_\infty^2 < \frac{|b|}{a}$$

and thus

$$ad < (a|c|)t_\infty^2 < |bc|.$$

This inequality is soluble only if $ad - bc = -1$. Similarly, in the case that $\left(\frac{-1, at_\infty^2+b}{\mathbb{R}}\right) = +1$ we have $at_\infty^2 + b > 0$ and thus $ct_\infty^2 + d > 0$. We deduce the Hilbert symbol condition is equivalent to

$$|bc| < (a|c|)t_\infty^2 < ad,$$

which is only soluble when $ad - bc = 1$.

□

These two lemmas together tell us that the source of the Brauer–Manin obstruction will come from the 2-adic properties of the surface $X_{a,b,c,d}$. The 2-adic solubility was spelled out by la Bréteche and Browning [13, Lemmas 4.9-4.15] and amounts to ensuring that (a, b, c, d) lie in certain specified congruence classes mod 16 (more details on this appear in the final section). Suppose that

$$\beta = v_2(b),\ \gamma = v_2(c) \text{ and } \delta = v_2(d),$$

are the 2-adic valuations of the coefficients and $\sigma(b, c, d) = (\epsilon_2, \epsilon_3, \epsilon_4)$. We make the change of variables

$$a = a',\ b = \epsilon_2 2^\beta b',\ c = \epsilon_3 2^\gamma c',\ d = \epsilon_4 2^\delta d'. \tag{4.3.1}$$

We will denote by $H^{\pm}_{\beta,\gamma,\delta}$ the union of congruence classes such that for $ad - bc = \pm 1$ we have

$$X_{a,b,c,d}(\mathbb{Q}_2) \neq \varnothing \iff (a', b', c', d') \in H^{\pm}_{\beta,\gamma,\delta} \bmod 16,$$

and similarly $\widetilde{H}^{\pm}_{\beta,\gamma,\delta}$ the union of congruence classes such that

$$X_{a,b,c,d}(\mathbb{Q}_2) \neq \varnothing \text{ but } X_{a,b,c,d}(\mathbb{Q}) = \varnothing \iff (a', b', c', d') \in \widetilde{H}^{\pm}_{\beta,\gamma,\delta} \bmod 16.$$

Our asymptotic formulae will be produced by counting points lying in residue classes in $H^{\pm}_{\beta,\gamma,\delta}$ or $\widetilde{H}^{\pm}_{\beta,\gamma,\delta}$, respectively, mod 16.

4.4 Calculating the asymptotics

We turn our attention to estimating the counting functions $N(B), N_{\mathrm{loc}}(B)$ and $N_{\mathrm{Br}}(B)$. We have seen in the preceeding sections that in order to do this we must count the number of points in S_{tot} that lie in particular congruence classes mod 16. To achieve this we appeal to a result based on homogeneous dynamics (Lemma 4.4.2), which allows us to count points on the quadric $ad - bc = \pm 1$ satisfying a congruence condition. We start by computing $N(B)$, the method for $N_{\mathrm{loc}}(B)$ and $N_{\mathrm{Br}}(B)$ being similar.

Theorem 4.4.1. *Let* $N(B) := \frac{1}{4}\#\{(a,b,c,d) \in S_{tot} : |(a,b,c,d)| \leqslant B\}$. *Then there exists* $\theta > 0$ *such that*

$$N(B) = \frac{32}{\pi^2}B^2 + O(B^{2-\theta}).$$

Recall that we had defined

$$S_{\mathrm{tot}}^{(1)} = \{(a,b,c,d) \in S_{\mathrm{tot}} : 2 \nmid a\},$$
$$S_{\mathrm{tot}}^{(2)} = \{(a,b,c,d) \in S_{\mathrm{tot}} : 2 \nmid a, 2 \mid b\}.$$

To count the elements in S_{tot}, we first denote the total set of allowable signs by S, the possible values of (β,γ,δ) for elements of $S_{\mathrm{tot}}^{(i)}$ by $L^{(i)}$ and the possible congruence classes elements can lie in by $T^{\pm}_{\beta,\gamma,\delta}$. Recall from Section 4.2 that there is a 1:4 correspondence between $\mathcal{M}$ and S_{tot}. Explicitly, we denote

$$S = \{(+,+,+,+),(+,-,+,-),(+,-,-,+),(+,+,-,-)\},$$
$$L^{(i)} = \{(\beta,\gamma,\delta) \in \mathbb{Z}^3_{\geqslant 0} : \min\{\beta+\gamma,\delta\} = 0 < \max\{\beta+\gamma,\delta\}, \beta \geqslant i-1\},$$
$$T^{\pm}_{\beta,\gamma,\delta} = \{\boldsymbol{\xi} \in \left((\mathbb{Z}/16\mathbb{Z})^{\times}\right)^4 : \epsilon_4 2^{\delta}\xi_1\xi_4 - \epsilon_2\epsilon_3 2^{\beta+\gamma}\xi_2\xi_3 \equiv \pm 1 \bmod 16\}.$$

Now we make the change of variables described in (4.3.1) so that our counting problem becomes

$$N(B) = \frac{1}{4}\sum_{\boldsymbol{\epsilon}\in S}\sum_{i\in\{1,2\}}\sum_{(\beta,\gamma,\delta)\in L^{(i)}}\sum_{\boldsymbol{\xi}\in T^{\pm}_{\beta,\gamma,\delta}}\left(K^{+}+K^{-}\right), \tag{4.4.1}$$

where $K^{\pm}$ denotes the total number of $(a,b',c',d') \in \mathbb{N}^4$ satisfying:

(i) $|(a, 2^{\beta}b', 2^{\gamma}c', 2^{\delta}d')| \leqslant B$,

(ii) $\epsilon_4 2^{\delta}ad' - \epsilon_2\epsilon_3 2^{\beta+\gamma}b'c' = \pm 1$,

(iii) $(a,b',c',d') \equiv \boldsymbol{\xi} \bmod 16$.

The following result of Browning and Gorodnik [16] provides an asymptotic formula for $K^{\pm}$, the main term of which factors as a product of local densities. The values of these densities are computed in Lemmas 4.4.3,4.4.4 and 4.4.5.

Lemma 4.4.2. *Let $F(\mathbf{x}) := \epsilon_4 2^\delta x_1 x_2 - \epsilon_2\epsilon_3 2^{\beta+\gamma} x_3 x_4$, then there exists $\theta > 0$ such that*

$$K^\pm = \mu_\infty^\pm(B) \prod_p \mu_p^\pm(\boldsymbol{\xi}) + O\left(\mu_\infty^\pm(B)^{1-\frac{\theta}{2}}\right),$$

where

$$\mu_p^\pm(\boldsymbol{\xi}) = \lim_{t\to\infty} p^{-3t} \#\{\mathbf{x} \in (\mathbb{Z}/p^t\mathbb{Z})^4 : F(\mathbf{x}) \equiv \pm 1 \bmod p^t,\ \mathbf{x} \equiv \boldsymbol{\xi} \bmod p^{v_p(16)}\}$$

$$\mu_\infty^\pm(B) = \lim_{\eta\to 0} \frac{1}{\eta} \int\limits_{\substack{0<x_1\leqslant B \\ 0<\epsilon_2 2^\beta x_2 \leqslant B \\ 0<\epsilon_3 2^\gamma x_3 \leqslant B \\ 0<\epsilon_4 2^\delta x_4 \leqslant B \\ |F(\mathbf{x})\mp 1|<\frac{\eta}{2}}} \mathrm{d}\mathbf{x}.$$

Proof. This is [16, Proposition 3.1] applied to the variety in which we are interested. □

We now proceed to compute these local density factors. The most important consequences of the following results are that $\mu_p^\pm$ is independent of $\boldsymbol{\xi}$ and that $\mu_\nu^+ = \mu_\nu^-$ for all $\nu \in \Omega$.

Lemma 4.4.3. *For p odd and any $\boldsymbol{\xi} \in (\mathbb{Z}/16\mathbb{Z})^\times$, we have $\mu_p^\pm(\boldsymbol{\xi}) = 1 - \frac{1}{p^2}$.*

Proof. For $p > 2$, under an obvious change of variables, the density is

$$\lim_{t\to\infty} p^{-3t} \#\{\mathbf{x} \in (\mathbb{Z}/p^t\mathbb{Z})^4 : x_1x_2 - x_3x_4 \equiv \pm 1 \bmod p^t\}.$$

Let $N(p^t)$ be the cardinality in which we are interested. Then we can express this congruence counting problem as an exponential sum.

$$\begin{aligned} N(p^t) &= p^{-t} \sum_{r,x_1,\ldots,x_4 \bmod p^t} e\left(\frac{(x_1x_2 - x_3x_4 \mp 1)r}{p^t}\right) \\ &= p^{-t} \sum_{r \bmod p^t} e\left(\frac{\mp r}{p^t}\right) \left| \sum_{x_1,x_2 \bmod p^t} e\left(\frac{r x_1 x_2}{p^t}\right)\right|^2. \end{aligned}$$

The sum over x_1, x_2 is equal to $p^t(r, p^t)$. Therefore,

$$\begin{aligned} N(p^t) &= p^t \sum_{r \bmod p^t} e\left(\frac{\mp r}{p^t}\right)(r,p^t)^2 \\ &= p^t \sum_{0\leqslant\alpha\leqslant t} p^{2\alpha} \sum_{\substack{r \bmod p^t \\ (r,p^t)=p^\alpha}} e\left(\frac{\mp r}{p^t}\right) \\ &= p^t \sum_{0\leqslant\alpha\leqslant t-1} p^{2\alpha} \sideset{}{^*}\sum_{r \bmod p^{t-\alpha}} e\left(\frac{\mp r}{p^{t-\alpha}}\right) + p^{3t}. \end{aligned}$$

This inner sum is now in the form of the Ramanujan sum $c_{p^{t-\alpha}}(\mp 1)$ so can be explicitly evaluated to give us

$$N(p^t) = p^{3t} - p^t \cdot p^{2(t-1)} = p^{3t}\left(1 - p^{-2}\right),$$

from which the result follows. □

Lemma 4.4.4. *We have*

$$\mu_\infty^{\pm}(B) = \frac{2B^2}{2^{\beta+\gamma+\delta}} + O(\log^2 B).$$

Proof. Under the change of variables

$$y_1 = x_1/B,\ y_2 = \epsilon_2 2^\beta x_2/B,\ y_3 = \epsilon_3 2^\gamma x_3/B,\ y_4 = \epsilon_4 2^\delta x_4/B \text{ and } \lambda = \eta/B^2,$$

we have

$$\mu_\infty^{\pm}(B) = \frac{\epsilon_2\epsilon_3\epsilon_4 B^2}{2^{\beta+\gamma+\delta}} \lim_{\lambda\to 0}\frac{1}{\lambda} \int\limits_{\substack{0<x_i\leqslant 1\\ |x_1x_4 - x_2x_3 \mp 1/B^2|<\lambda/2}} \mathrm{d}\mathbf{x}.$$

Recall the set of allowable signs

$$S = \{(+,+,+,+),(+,-,+,-),(+,-,-,+),(+,+,-,-)\},$$

and observe that $\epsilon_2\epsilon_3\epsilon_4 = 1$. We now change to hyperbolic co-ordinates

$$\begin{aligned} u_1 &= \sqrt{x_1x_4},\ v_1 = \log\left(\sqrt{\tfrac{x_1}{x_4}}\right),\\ u_2 &= \sqrt{x_2x_3},\ v_2 = \log\left(\sqrt{\tfrac{x_2}{x_3}}\right). \end{aligned}$$

The integral above transforms to $16I$ where

$$\begin{aligned} I :&= \frac{1}{4} \int\limits_{\substack{0<u_i\leqslant 1\\ |u_1^2-u_2^2\mp 1/B^2|<\lambda/2}} u_1u_2 \int\limits_{-\log u_i<v_i<\log u_i} \mathrm{d}v_1\mathrm{d}v_2\mathrm{d}u_1\mathrm{d}u_2\\ &= \int\limits_{\substack{0<u_i\leqslant 1\\ |u_1^2-u_2^2\mp 1/B^2|<\lambda/2}} (u_1\log u_1)(u_2\log u_2)\mathrm{d}u_1\mathrm{d}u_2. \end{aligned}$$

At this point, we observe that one can change the sign of the $\pm$ symbol in the expression $|u_1^2 - u_2^2 \mp \frac{1}{B^2}| < \frac{\lambda}{2}$ by interchanging the variables u_2 and u_1. Therefore we need only restrict our attention to the case $|u_1^2 - u_2^2 - \frac{1}{B^2}| < \frac{\lambda}{2}$. Noting that we may take λ small enough that

$$\frac{1}{B^2} - \frac{\lambda}{2} > 0,$$

we see that the region of integration is the space between the two hyperbolae $u_1^2 - u_2^2 = \frac{1}{B^2} + \frac{\lambda}{2}$ and $u_1^2 - u_2^2 = \frac{1}{B^2} - \frac{\lambda}{2}$ inside the box $[0,1]^2$. The major

contribution occurs before the lower hyperbola meets the line $u_1 = 1$. The rest of the integral will contribute $o(\lambda)$. With this in mind write

$$I = I_1 + I_2,$$

where

$$I_1 = \int_0^{\sqrt{1-1/B^2-\lambda/2}} u_2 \log u_2 \int_{\sqrt{u_2^2+1/B^2-\lambda/2}}^{\sqrt{u_2^2+1/B^2+\lambda/2}} u_1 \log u_1 \mathrm{d}u_1 \mathrm{d}u_2$$

$$I_2 = \int_{\sqrt{1-1/B^2-\lambda/2}}^{\sqrt{1-1/B^2+\lambda/2}} u_2 \log u_2 \int_{\sqrt{u_2^2+1/B^2-\lambda/2}}^{1} u_1 \log u_1 \mathrm{d}u_1 \mathrm{d}u_2.$$

The result of the inner integral of I_1 is

$$\frac{1}{4}\left[\left(c+\frac{\lambda}{2}\right)\left(\log\left(c+\frac{\lambda}{2}\right)-1\right)-\left(c-\frac{\lambda}{2}\right)\left(\log\left(c-\frac{\lambda}{2}\right)-1\right)\right], \quad (4.4.2)$$

where

$$c = u_2^2 + \frac{1}{B^2}.$$

By considering the Taylor series of $\log(c \pm \frac{\lambda}{2})$, we write this expression as

$$I_1 = \frac{\lambda}{4} \int_0^{\sqrt{1-1/B^2-\lambda/2}} u_2 \log u_2 \log\left(u_2^2 + \frac{1}{B^2}\right) \mathrm{d}u_2 + O\left(\lambda^3\right).$$

We now split the range of this integral according to whether $|u^2B^2| < 1$ or $|u_2^2B^2| > 1$. In the first of these ranges, we have

$$\begin{aligned}\int_0^{1/B} u_2 \log u_2 \log(u_2^2 + \frac{1}{B^2})\mathrm{d}u_2 &= \int_0^{1/B} u_2 \log u_2 \left[-2\log B + O(u_2^2B^2)\right] \mathrm{d}u_2 \\ &\ll \frac{\log^2 B}{B^2}.\end{aligned}$$

Meanwhile, in the second range,

$$\begin{aligned}&\int_{1/B}^{\sqrt{1-1/B^2-\lambda/2}} u_2 \log u_2 \log(u_2^2 + \frac{1}{B^2})\mathrm{d}u_2 \\ &\qquad = \int_{1/B}^{\sqrt{1-1/B^2-\lambda/2}} u_2 \log u_2 \left[2\log u_2 + O(B^{-2}u_2^{-2})\right] \mathrm{d}u_2 \\ &\qquad = 2\int_{1/B}^{\sqrt{1-1/B^2-\lambda/2}} u_2 \log^2 u_2 \mathrm{d}u_2 + O(\frac{\log B}{B^2}) \\ &\qquad = \frac{1}{2} + O(\frac{\log^2 B}{B^2}).\end{aligned}$$

We now turn to estimating I_2 by bounding above trivially then using the binomial theorem to see

$$\begin{aligned} I_2 &\leqslant \left(\sqrt{1-\frac{1}{B^2}+\frac{\lambda}{2}}-\sqrt{1-\frac{1}{B^2}-\frac{\lambda}{2}}\right)\left(1-\sqrt{1-\lambda}\right)e^{-2} \\ &= e^{-2}\left(\frac{1}{2}\lambda+\frac{1}{4}\frac{\lambda}{B^2}+\dots\right)(1-\sqrt{1-\lambda}), \end{aligned}$$

which clearly tends to 0 after dividing by λ and taking the limit $\lambda \to 0$. □

Lemma 4.4.5. *Let $(\beta,\gamma,\delta) \in L^{(i)}$. For any choice of $\boldsymbol{\xi} \in T^{\pm}_{\beta,\gamma,\delta}$, we have*

$$\mu_2^{\pm}(\boldsymbol{\xi}) = 2^{-12}.$$

Proof. Denote

$$S(t) := \{\mathbf{x} \in \left(\mathbb{Z}/2^t\mathbb{Z}\right)^4 : F(\mathbf{x}) \equiv \pm 1 \bmod 2^t \text{ and } \mathbf{x} \equiv \boldsymbol{\xi} \bmod 16\},$$

for $t \geqslant 4$ so that

$$\mu_2(B) = \lim_{t\to\infty} 2^{-3t}\#S(t).$$

We'll use Hensel's lemma to relate this to $S(4) = \{\boldsymbol{\xi}\}$. In order to do this we first, note that

$$v_2(\nabla F(\mathbf{x})) = \min\{\delta + v_2(x_4), \beta+\gamma+v_2(x_3), \beta+\gamma+v_2(x_2), \delta+v_2(x_1)\}.$$

If $\mathbf{x} \in S(t)$ for any $t \geqslant 4$ then $x_i \equiv \xi_i \bmod 16$ and therefore $v_2(x_i) = 0$ for $i = 1,\dots,4$. Moreover either $\delta = 0$ or $\beta+\gamma = 0$, since $(\beta,\gamma,\delta) \in L$. Hence

$$v_2(\nabla F(\mathbf{x})) = 0.$$

Thus by Hensel's lemma, we have

$$\#S(t) = 2^3\#S(t-1) = \dots = 2^{3(t-4)}\#S(4) = 2^{3(t-4)},$$

which completes the proof. □

Combining these results with expression (4.4.1) and writing

$$T_{\beta,\gamma,\delta} = T^{+}_{\beta,\gamma,\delta} \sqcup T^{-}_{\beta,\gamma,\delta},$$

we see that

$$N(B) = \frac{B^2}{2^8\pi^2}\sum_{i\in\{1,2\}}\sum_{(\beta,\gamma,\delta)\in L^{(i)}}\frac{\#T_{\beta,\gamma,\delta}}{2^{\beta+\gamma+\delta}} + O(B^{2-\theta}). \tag{4.4.3}$$

This inner sum is easily verified to be 2^{13} completing the proof of Theorem 4.4.1.

The other counting functions can be evaluated similarly, using Lemma 4.4.2 again although with a slightly different setup. We will make the same change

of variables as before although this time the set of allowable signs and congruence classes will be informed by Section 4.3. In particular, (4.1.1) clearly has no real solutions if $\sigma(b,c,d) = (+,-,-)$, so we exclude this possibility. Then, analogously to above our local solution counting function can be expressed as

$$N_{\mathrm{loc}}(B) = \frac{1}{4} \sum_{\boldsymbol{\epsilon} \in S \setminus \{(+,-,-)\}} \sum_{i \in \{1,2\}} \sum_{(\beta,\gamma,\delta) \in L^{(i)}} \sum_{\boldsymbol{\xi} \in H^{\pm}_{\beta,\gamma,\delta}} \left(K^{+} + K^{-}\right),$$

where $H_{\beta,\gamma,\delta} = H^{+}_{\beta,\gamma,\delta} \cup H^{-}_{\beta,\gamma,\delta}$. The $K^{\pm}$ appearing in this expression is exactly the same as in (4.4.1) so we may use Lemma 4.4.2 to write this as

$$N_{\mathrm{loc}}(B) = \frac{B^2}{2^{10}\pi^2} \sum_{\boldsymbol{\epsilon} \in S \setminus \{(+,-,-)\}} \sum_{i \in \{1,2\}} \sum_{(\beta,\gamma,\delta) \in L^{(i)}} \frac{\# H_{\beta,\gamma,\delta}}{2^{\beta+\gamma+\delta}} + O(B^{2-\theta}). \quad (4.4.4)$$

The conditions describing $H_{\beta,\gamma,\delta}$ are spelled out explicitly by Lemmas 4.8-4.15 of [13]. Crucially, the size of $H_{\beta,\gamma,\delta}$ does not depend on $\boldsymbol{\epsilon}$, so one can compute the sum

$$\tau_{loc,2} := 2^{-13} \sum_{i \in \{1,2\}} \sum_{(\beta,\gamma,\delta) \in L^{(i)}} \frac{\# H_{\beta,\gamma,\delta}}{2^{\beta+\gamma+\delta}} > 0 \quad (4.4.5)$$

to obtain the expression (4.1.3).

The evaluation of $N_{\mathrm{Br}}(B)$ will follow similar lines except this time counting over different congruence classes mod 16, again described explicitly in [13] and with a smaller set of allowable signs. Recall from Section 4.3 that whenever $\sigma(b,c,d) = (+,-,+)$ the surface $X_{a,b,c,d}$ always satisfies the Hasse principle. Therefore analogously to (4.4.4), we have

$$N_{\mathrm{Br}}(B) = \frac{B^2}{2^{10}\pi^2} \sum_{\boldsymbol{\epsilon} \in S \setminus \{(+,-,-),(-,+,-)\}} \sum_{i \in \{1,2\}} \sum_{(\beta,\gamma,\delta) \in L^{(i)}} \frac{\# \widetilde{H}_{\beta,\gamma,\delta}}{2^{\beta+\gamma+\delta}} + O(B^{2-\theta}). \quad (4.4.6)$$

where

$$0 < 2^{-13} \sum_{i \in \{1,2\}} \sum_{(\beta,\gamma,\delta) \in L^{(i)}} \frac{\# \widetilde{H}_{\beta,\gamma,\delta}}{2^{\beta+\gamma+\delta}} =: \sigma_{loc,2} \leqslant \tau_{loc,2}. \quad (4.4.7)$$

This completes the proof of Theorem 4.1.1.

All that remains is to give a value to the 2-adic densities $\tau_{loc,2}, \sigma_{loc,2}$ defined above. In Table 4.4, we record the values for $\# H_{\beta,\gamma,\delta}$ and $\# \widetilde{H}_{\beta,\gamma,\delta}$ which were obtained by using MATLAB to enumerate the residue classes in $((\mathbb{Z}/16\mathbb{Z})^{\times})^4$ which satisfy the conditions outlined in [13]. A more detailed discussion of the la Bretèche–Browning conditions and the computation of the size of these sets of congruence classes is relagated to the final section to this chapter.
With these values, it is simple to conclude from (4.4.5) that

$$\tau_{loc,2} = \frac{17856}{3 \times 2^{13}}.$$

Table 4.1: Computing the size of $H_{\beta,\gamma,\delta}$ and $\widetilde{H}_{\beta,\gamma,\delta}$.

$\boldsymbol{\beta}$	$\boldsymbol{\gamma}$	$\boldsymbol{\delta}$	$\#\boldsymbol{H_{\beta,\gamma,\delta}}$	$\#\boldsymbol{\widetilde{H}_{\beta,\gamma,\delta}}$
0	0	1	1024	416
0	0	$\geqslant$2 even	960	512
0	0	$\geqslant$3 odd	1024	544
0	1	0	1024	416
0	$\geqslant$2 even	0	960	448
0	$\geqslant$3 odd	0	1024	416
1	0	0	1024	192
1	1	0	1024	128
1	$\geqslant$2 even	0	1024	320
1	$\geqslant$3 odd	0	1024	256
2	0	0	992	480
2	1	0	1024	576
2	$\geqslant$2	0	768	320
3	0	0	1024	320
3	1	0	1024	384
3	$\geqslant$2	0	768	128
$\geqslant$4 even	0	0	992	480
$\geqslant$4 even	1	0	1024	576
$\geqslant$4 even	$\geqslant$2	0	768	320
$\geqslant$4 odd	0	0	1024	320
$\geqslant$4 odd	1	0	1024	384
$\geqslant$4 odd	$\geqslant$2	0	768	128

Likewise, we compute

$$\sigma_{loc,2} = \frac{2112}{2^{13}}.$$

4.5 2-adic density calculations

This section is concerned with proving the values claimed for the 2-adic densities $\tau_{loc,2}$ and $\sigma_{loc,2}$. In order to do this we quote the local conditions developed by la Bretèche–Browning. Denote $\mathscr{D} := \{2^n(1+4m) : m \in \mathbb{Z}_2, n \in \mathbb{Z}\}$ and $\overline{\mathscr{D}} := \{2^n(3+4m) : m \in \mathbb{Z}_2, n \in \mathbb{Z}\}$. Let

$$T_{\mathrm{tot}} := \{t = (A,B,C,D) \in \mathbb{Z}^4 : 2 \nmid A, AD - BC = 1\}.$$

Write $B = 2^\beta B', C = 2^\gamma C'$ and $D = 2^\delta D'$. In [13], the authors denote by $T^{(2)}_{\mathrm{tot}}$ those elements in T_{tot} for which the associated Châtelet surface has $\mathbb{Q}_2$ points. They also introduce $T^{(1)}_{\mathrm{tot}}$ which in their analysis detects when a certain torsor has $\mathbb{Q}_2$ points. This torsor is defined by the equations

$$\begin{cases} 0 \neq Y_1^2 + Z_1^2 = e_1(aU^2 + bV^2) \\ 0 \neq Y_2^2 + Z_2^2 = e_2(cU^2 + dV^2), \end{cases}$$

for some suitable choice of e_1, e_2 which are both squarefree numbers satisfying $p \mid e_i \implies p \equiv 3 \bmod 4$. The set $T^{(1)}_{\mathrm{tot}}$ is defined to be those elements of T_{tot} for which both $aU^2 + BV^2$ and $cU^2 + dV^2$ are in $\mathscr{D}$. This corresponds exactly to the local Hilbert symbol taking the value $+1$. We wish to understand when $(a,b,c,d) \in T^{(2)}_{\mathrm{tot}}$ and when $(a,b,c,d) \in T^{(2)}_{\mathrm{tot}} \setminus T^{(1)}_{\mathrm{tot}}$. The conditions for this will take the form of a union of congruence classes modulo 16. The important fact about this set of classes is that their cardinality is independent of the sign of the coefficients and of the residue classes of the e_i. If $\sigma(b,c,d) = (-,-,+)$ we saw in Lemma 4.3.2 that the Hilbert symbol at the real place took the value $ad - bc$ and so to find a Hasse failure we need the 2-adic symbol to be $bc - ad$. However the invariance of the size of $H_{\beta,\gamma,\delta}$ and $\widetilde{H}_{\beta,\gamma,\delta}$ with respect to the sign means we need not worry about treating this as a seperate case.

Below we reproduce the list of lemmas in [13] which determine the sets $T^{(1)}_{\mathrm{tot}}$ and $T^{(2)}_{\mathrm{tot}}$. The lemmas are seperated depending on the congruence class of A and B. Let $T^{(k)}_{\mathrm{tot}}(i,j)$ denote the elements of $T^{(k)}_{\mathrm{tot}}$ such that $(A,B) \equiv (i,j) \bmod 4$. We will repeatedly use

$$T^{(2)}_{\mathrm{tot}}(i,j) = -T^{(2)}_{\mathrm{tot}}(-i,-j). \tag{4.5.1}$$

Lemma 4.5.1. *We have $t \in T^{(1)}_{tot}(1,1)$ if and only if one of the following holds:*

- *$C \in \mathscr{D}$ or $D \in \mathscr{D}$;*
- *$(C,D) \equiv (3,3) \bmod 4$ with $A + B \equiv C + D \equiv 2 \bmod 8$;*
- *$C \equiv 3 \bmod 4$, $\delta \geqslant 1$ and $D' \equiv 3 \bmod 4$ such that $C + D' \equiv 2 \bmod 8$ if $2 \mid \delta$;*

- $D \equiv 3 \bmod 4$, $\gamma \geqslant 1$ *and* $C' \equiv 3 \bmod 4$ *such that* $C' + D \equiv 2 \bmod 8$ *if* $2 \mid \gamma$.

Moreover $t \in T_{tot}^{(2)}(1,1)$ *if and only if one of the following holds:*

- $C \in \mathscr{D}$ *or* $D \in \mathscr{D}$*;*
- $(C, D) \equiv (3,3) \bmod 4$ *with* $A + B \equiv C + D \equiv 2 \bmod 8$*;*
- $C \equiv 3 \bmod 4$ *and* $D' \equiv 3 \bmod 4$ *with one of the following:*
  - $2 \nmid \delta$,
  - $2 \mid \delta$, $\delta \geqslant 2$ *and* $C + D' \equiv 2 \bmod 8$,
  - $2 \mid \delta$, $\delta \geqslant 2$ *and* $A + B \equiv C + D' \equiv 6 \bmod 8$.
- $D \equiv 3 \bmod 4$ *and* $C' \equiv 3 \bmod 4$ *with one of the following:*
  - $2 \nmid \gamma$,
  - $2 \mid \gamma$, $\gamma \geqslant 2$ *and* $C' + D \equiv 2 \bmod 8$,
  - $2 \mid \gamma$, $\gamma \geqslant 2$ *and* $A + B \equiv C' + D \equiv 6 \bmod 8$.

Lemma 4.5.2. *We have* $t \in T_{tot}^{(1)}(3,3)$ *if an only if:*

- $A + B \equiv 2 \bmod 8$ *and* $C + D \in \{0,1,2,4,5\} \bmod 8$.

Moreover the characterisation of $T_{tot}^{(2)}(3,3)$ *follows from* (4.5.1) *and Lemma 4.5.1.*

Lemma 4.5.3. *We have* $t \in T_{tot}^{(1)}(1,3)$ *if and only if one of the following*

- $C \in \mathscr{D}$*;*
- $C \in \overline{\mathscr{D}}$ *and* $C + D \in \{1,2,5\} \bmod 8$*;*
- $D \equiv 1 \bmod 4$, $\gamma = 1$ *and* $C' \equiv 3 \bmod 4$*;*
- $D \equiv 3 \bmod 4$, $\gamma \geqslant 2$ *and* $C' \equiv 3 \bmod 4$ *with* $C' + D \equiv 2 \bmod 8$ *if* $2 \mid \gamma$*;*
- $(C, D) \equiv (3,3) \bmod 4$ *with one of the following:*
  - $A + B \equiv 0 \bmod 8$,
  - $A + B \equiv 4 \bmod 8$ *and* $C + D \equiv 0 \bmod 8$,
  - $A + B \equiv 4 \bmod 8$ *and* $A + B = C + D \bmod 16$.

Moreover $t \in T_{tot}^{(2)}(1,3)$ *if and only if one of the following holds:*

- $C \in \mathscr{D}$*;*
- $C \in \overline{\mathscr{D}}$, $2 \mid C$*;*
- $C \equiv 3 \bmod 4$ *and* $C + D \in \{1,2,3,5,6,7\} \bmod 8$*;*

- $(C, D) \equiv (3, 1) \bmod 4$, *with one of the following:*
  - $A + B \equiv 0 \bmod 8$,
  - $A + B \equiv 4 \bmod 8$ *and* $C + D \equiv 0 \bmod 8$,
  - $A + B \equiv 4 \bmod 8$ *and* $A + B \equiv C + D \bmod 16$.

Lemma 4.5.4. *We have* $t \in T^{(1)}_{tot}(1, 0)$ *if and only if one of the following holds:*

- $C \in \mathscr{D}$ *or* $C + D \in \{0, 1, 2, 4, 5\} \bmod 8$ *with* $4 \nmid D$;
- $\gamma = 1$ *and* $(C', D) \equiv (3, 1) \bmod 4$;
- $\gamma \geqslant 2$ *and* $(C', D) \equiv (3, 3) \bmod 4$, *with* $C' + D \equiv 2 \bmod 8$ *if* $2 \mid \gamma$.

Moreover $t \in T^{(2)}_{tot}(1, 0)$ *if and only if one of the following holds:*

- $C \in \mathscr{D}$ *and* $4 \nmid D$;
- $C \in \mathscr{D}$ *and* $C + D \in \{0, 1, 2, 4, 5\} \bmod 8$;
- $C \in \mathscr{D}$, $4 \nmid D$ *and* $C + D \in \{3, 6, 7\} \bmod 8$, *with one of the following:*
  - $C \equiv 3 \bmod 4$ *and* $C + D \equiv 6 \bmod 8$, *with* $A + B' \equiv 6 \bmod 8$ *if* $B' \equiv 1 \bmod 4$ *and* $2 \mid \beta$;
  - $C' \equiv 3 \bmod 4$, $\gamma = 1$ *and* $D \equiv 1 \bmod 4$;
  - $C' \equiv 3 \bmod 4$, $\gamma \geqslant 2$ *and* $D \equiv 3 \bmod 4$, *with one of the following:*
    - $2 \nmid \gamma$,
    - $2 \mid \gamma$ *and* $C' + D \equiv 2 \bmod 8$,
    - $2 \mid \gamma$ *and* $C' + D \equiv 6 \bmod 8$, *with* $A + B' \equiv 6 \bmod 8$ *if* $B' \equiv 1 \bmod 4$ *and* $2 \mid \beta$.

Lemma 4.5.5. *We have* $T^{(1)}_{tot}(3, 0)$ *if and only if one of the following holds:*

- $D \in \mathscr{D}$ *and* $4 \nmid D$, *with* $A + B' \equiv 2 \bmod 8$ *if* $2 \mid \beta$;
- $\delta = 1$ *and* $D' \equiv 3 \bmod 4$ *with* $\beta \in \{2, 3\}$.

Moreover the characterisation of $T^{(2)}_{tot}(3, 0)$ *follows* (4.5.1) *and Lemma 4.5.4.*

Lemma 4.5.6. *We have* $t \in T^{(1)}_{tot}(1, 2)$ *if and only if one of the following holds:*

- $C \in \mathscr{D}$;
- $C' \equiv 3 \bmod 4$, *with one of the following:*
  - $D \equiv 1 \bmod 4$, $\gamma \geqslant 1$,
  - $D \equiv 3 \bmod 4$, $\gamma \geqslant 1$ *and* $C' + D \equiv 2 \bmod 8$ *if* $2 \mid \gamma$,
  - $B' \equiv 3 \bmod 4$, $\gamma = 0$, $\delta \in \{0, 2, 3\}$ *and* $D' \equiv 1 \bmod 4$,

- $B' \equiv 3 \bmod 4$, $\gamma = 0$, $\delta \in \{1,2,3\}$ *and* $D' \equiv 3 \bmod 4$, *with* $C + D' \equiv 2 \bmod 8$ *if* $\delta = 2$,
- $B' \equiv 1 \bmod 4$, $\gamma = 0$ *and* $D' \equiv 1 \bmod 4$,
- $B' \equiv 1 \bmod 4$, $\gamma = 0$, $\delta \geqslant 3$ *and* $D' \equiv 3 \bmod 4$, *with* $C + D' \equiv 2 \bmod 8$ *if* $2 \mid \delta$.

Moreover $t \in T_{tot}^{(2)}(1,2)$ *if and only if one of the following holds:*

- $C \in \mathscr{D}$;
- $C' \equiv 3 \bmod 4$, $\gamma \geqslant 1$;
- $C \equiv 3 \bmod 4$, $\gamma = 0$, *with one of the following:*
  - $B' \equiv D' \bmod 4$,
  - $(B', D') \in \{(3,1),(1,3)\} \bmod 4$, $\delta \neq 1$.

Lemma 4.5.7. *We have* $t \in T_{tot}^{(1)}(3,2)$ *if and only if one of the following holds:*

- $C + D \in \{0,1,2,4,5\} \bmod 8$;
- $D \equiv 1 \bmod 4$ *and* $C + D \in \{3,6\} \bmod 8$;
- $C + D \equiv 3 \bmod 8$, $\delta = 1$ *and* $D' \equiv B' \bmod 4$;
- $C \equiv 3 \bmod 4$, $\delta \geqslant 2$, *with one of the following:*
  - $B' \equiv 3 \bmod 4$, $\delta = 3$,
  - $B' \equiv 3 \bmod 4$, $\delta = 2$ *and* $C + D' \in \{0,2,4\} \bmod 8$,
  - $(B', D') \equiv (1,1) \bmod 4$,
  - $(B', D') \equiv (1,3) \bmod 4$, $\delta \geqslant 3$ *and* $C + D' \equiv 2 \bmod 8$ *if* $2 \mid \delta$.

Moreover the characterisation of $T_{tot}^{(2)}(3,2)$ *follows from* (4.5.1) *and Lemma 4.5.6.*

Now to compute the sums defining the local densities (4.4.5) and (4.4.7), we must sum over all possible residue classes $(A,B,C,D) \bmod 16$ which the (a,b,c,d) can lie in under the conditions that $2^\beta \mid B, 2^\gamma \mid C, 2^\delta \mid D$. We enumerate how many of these satisfy the relevant conditions in the previous lemmas in Table 4.4. Finally we divide by the factor $2^{\beta+\gamma+\delta}$ and then compute the sum over all possible values of β, γ and δ.

Chapter 5

A sieve for points on a hypersurface

The purpose of this chapter is to produce a large sieve style upper bound for points on a hypersurface (or a product of hypersurfaces) which lie in given residue classes. In Section 5.2 this will find several applications to the study of rational points, in particular the extension of the Loughran–Smeets upper bound to the setting of a hypersurface base (Theorem B in Chapter 1). The sieve theorem itself and the application to thin sets and rational curves (5.3.4 and 5.3.8) featured in the paper [63].

5.1 Main theorem

Let $F \in \mathbb{Z}[x_0, \dots, x_N]$ be a homogeneous form of degree d. Denote by X the projective variety defined by F and assume (for simplicity) that it is smooth. There is a natural model $\mathcal{X}$ given by the closure of X in $\mathbb{P}^N_{\mathbb{Z}}$ (i.e. the $\mathbb{Z}$-scheme defined by the same equation) for the duration of this chapter we will abuse notation and write $X(\mathbb{Z}) = \mathcal{X}(\mathbb{Z})$ and $X(\mathbb{Z}/m\mathbb{Z}) = \mathcal{X}(\mathbb{Z}/m\mathbb{Z})$. Fix $m, n \in \mathbb{N}$. Let

$$\Omega_{p^m} \subset \{(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}) \in (\mathbb{Z}/p^m\mathbb{Z})^{n(N+1)} : p \nmid \mathbf{x}^{(j)}, F(\mathbf{x}^{(j)}) \equiv 0 \bmod p^m \, \forall j\} \tag{5.1.1}$$

be some non-empty collection of residue classes for each prime p. Denote their relative density by

$$\omega_p := 1 - \frac{\#\Omega_{p^m}}{\#\widehat{X^n}\,(\mathbb{Z}/p^m\mathbb{Z})} \in [0, 1),$$

where $\widehat{X^n}$ denotes the affine cone of X^n. We will establish a bound for points on hypersurfaces of bounded height (see the end of this subsection for the description of the particular height we choose) lying in these prescribed residue classes.

Lemma 5.1.1. *Assume that $N > 2^d(d-1)-1$. Then there exist $\delta_1, \delta_2 > 0$ such that*

$$\#\{x \in X(\mathbb{Q}) : H(x) \leqslant B \text{ and } x \bmod p^m \in \Omega_{p^m} \forall p\} \ll \frac{B}{\min\{G(B^{\delta_1}), B^{\delta_2}\}},$$

where

$$G(Q) = \sum_{q \leqslant Q} \mu^2(q) \prod_{p|q} \frac{\omega_p}{1-\omega_p}.$$

(ω_p is as defined above for $n = 1$.)

This generalises the degree 2 case proven by Browning–Loughran [19, Thm 1.7]. Lemma 5.1.1 will be sufficient for most of our applications, however Corollary 5.3.8 requires a mild extension to powers of hypersurfaces. In this case we need to investigate the subset $\mathbf{N}(X^n, \mathbf{B}, \Omega)$ defined by

$$\#\{x \in X^n(\mathbb{Q}) : B_j/2 < ||\mathbf{x}^{(j)}||_\infty \leqslant B_j \text{ and } (\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}) \bmod p^m \in \Omega_{p^m} \forall p, j\},$$

where $(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)})$ is a representative of x in $(\mathbb{Z}^{N+1}_{\text{prim}})^n$. We have the following analogous estimate for products.

Theorem 5.1.2. *Suppose $N > 2^d(d-1)-1$. Then for any $\epsilon > 0$ and any $Q \geqslant 1$, one has*

$$\mathbf{N}(X^n, \mathbf{B}, \Omega) \ll_{\epsilon, X} (B_1 \dots B_n)^{N+1-d} \left(\frac{1}{G(Q)} + E_1 + E_2 \right),$$

where

$$E_1 = \min_j \{B_j\}^{-\frac{1}{2}} Q^{m(2d-2n+1)+2+\epsilon}$$

$$E_2 = \min_j \{B_j\}^{-\frac{(N+1)2^{-d}-(d-1)}{4d}+\epsilon} Q^{2-m(d+2n)+\frac{5md(N+1)}{2^{d-1}(d-1)}+\frac{(N+1)m2^{-d}-m(d-1)}{2d}-\epsilon}.$$

Note that Lemma 5.1.1 follows immediately from this.

Remark. The same proof would allow one to establish a large sieve result not just for powers of a hypersurface but also for products of the form $X_1 \times \cdots \times X_n$ where each X_i is a hypersurface defined by a form in $N_i + 1$ variables of degree d_i such that $N_i > 2^{d_i}(d_i - 1) - 1$.

Throughout this chapter, we count points of bounded height. The height that we choose is a standard anticanonical height function. If $X \subset \mathbb{P}^N$ is a hypersurface as defined above and $x \in X(\mathbb{Q})$ is such that $x = (x_0 : \cdots : x_N)$ where $(x_0, \dots, x_N) \in \mathbb{Z}^{N+1}_{\text{prim}}$ then

$$H(x) = ||(x_0, \dots, x_N)||^{N+1-d} =: \max_i |x_i|^{N+1-d}.$$

The height on a product of hypersurfaces is just the product of the heights.

5.2 Proof of Theorem 5.1.2

Since the degree 2 case is already known, we will restrict our attention to $d \geqslant 3$. We count points $x \in X^n(\mathbb{Q})$ via their representatives $\mathbf{x} = (\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}) \in \mathbb{Z}^{n(N+1)}$ where the $\mathbf{x}^{(j)}$ are primitive vectors with $F(\mathbf{x}^{(j)}) = 0$. Passing to the affine cone, we see that we may bound $\mathbf{N}(X^n, \mathbf{B}, \Omega)$ by

$$\#\{\mathbf{x} \in \left(\mathbb{Z}^{N+1}_{\text{prim}}\right)^n : B_j/2 < ||\mathbf{x}^{(j)}|| \leqslant B_j, F(\mathbf{x}^{(j)}) = 0 \text{ and } \mathbf{x} \bmod p^m \in \Omega_{p^m} \forall p\}.$$

This quantity can be bounded above using the Selberg sieve. Let

$$P(Q) = \prod_{\substack{p<Q \\ \omega_p>0 \\ p\nmid \operatorname{disc}(F)}} p \quad \text{and} \quad \Lambda(\mathbf{x}) = \prod_{\substack{p|P(Q) \\ \mathbf{x} \bmod p^m \in \Omega^C_{p^m}}} p,$$

where $\Omega^C_{p^m} = \widehat{X^n}(\mathbb{Z}/p^m\mathbb{Z}) \setminus \Omega_{p^m}$.

Define a sequence $\mathcal{A} = (a_\lambda)$ of non-negative numbers, supported on finitely many integers λ, by

$$a_\lambda = \sum_{\substack{\mathbf{x} \in (\mathbb{Z}^{(N+1)}_{\text{prim}})^n \\ F(\mathbf{x}^{(j)})=0 \\ \Lambda(\mathbf{x})=\lambda}} \prod_{j=1}^{n} \prod_{i=0}^{N} W\left(\frac{x_i^{(j)}}{B_j}\right),$$

for W some appropriate smooth, compactly supported weight function. Then,

$$\sum_{(\lambda, P(Q))=1} a_\lambda = \sum_{\substack{\mathbf{x} \in (\mathbb{Z}^{(N+1)}_{\text{prim}})^n \\ F(\mathbf{x}^{(j)})=0 \\ (\Lambda(\mathbf{x}), P(Q))=1}} \prod_{j=1}^{n} \prod_{i=0}^{N} W\left(\frac{x_i^{(j)}}{B_j}\right) = \sum_{\substack{\mathbf{x} \in (\mathbb{Z}^{(N+1)}_{\text{prim}})^n \\ F(\mathbf{x}^{(j)})=0 \\ \mathbf{x} \bmod p^m \in \Omega_{p^m} \forall p | P(Q)}} \prod_{j=1}^{n} \prod_{i=0}^{N} W\left(\frac{x_i^{(j)}}{B_j}\right).$$

Theorem 5.1.2 will follow from a suitable upper bound for the a_λ sum. This is achieved by an appeal to Selberg's upper bound sieve.

Lemma 5.2.1 ([46, Theorem 7.1])**.** *Let $\mathcal{A} = (a_n)$ be a finite sequence of non-negative numbers and let P be a finite product of distinct primes. For every $d \mid P$ suppose that we have*

$$\sum_{n \equiv 0 \bmod d} a_n = g(d)X + r_d(\mathcal{A}),$$

where $X > 0$ and g is a multiplicative function with $0 < g(p) < 1$ for $p \mid P$. Then for any $D > 1$, we have

$$\sum_{\gcd(n,P)=1} a_n \leqslant X \left(\sum_{\substack{d|P \\ d \leqslant \sqrt{D}}} \frac{g(d)}{1-g(d)} \right)^{-1} + \sum_{\substack{d|P \\ d<D}} \tau_3(d) |r_d(\mathcal{A})|.$$

In order to apply this, we need an expression of the form

$$\sum_{\lambda \equiv 0 \bmod q} a_\lambda = g(q)Y + r_q(\mathcal{A}),$$

for a constant Y and suitable multiplicative function g and small remainder term $r_q(\mathcal{A})$. This information will be provided by the following result of Schindler–Sofos [102, Lemma 2.1].

Lemma 5.2.2. *Let $g \in \mathbb{Z}[x_0, \dots, x_N]$ a polynomial of degree $d \geqslant 3$. Fix $R > 0$ and $\mathbf{z} \in \mathbb{Z}^{N+1}$. If $N + 1 > 2^d(d-1)$, then one has*

$$\sum_{\substack{\mathbf{x} \in \mathbb{Z}^{N+1} \\ g(\mathbf{x})=0}} \prod_{i=0}^{N} W\Big(\frac{x_i}{R} - z_i\Big) - \mathfrak{S}\mathfrak{J}_W \ll_\epsilon R^{N-d+\frac{1}{2}} + ||g||^{\frac{5(N+1)}{2^d(d-1)} - \frac{3}{2}} R^{N+1-d+\epsilon - \frac{(N+1)2^{-d} - (d-1)}{4d}},$$

where

$$\mathfrak{J}_W = \int_{-\infty}^{\infty} \int_{\mathbb{R}^{N+1}} e(\gamma g(\mathbf{u})) \prod_{i=0}^{N} W\left(\frac{u_i}{R}\right) \mathrm{d}\mathbf{u}\mathrm{d}\gamma$$
$$\mathfrak{S} = \prod_p \sigma_p(g).$$

Here σ_p is the local density defined as

$$\sigma_p(g) := \lim_{\ell \to \infty} p^{-\ell N} \#\{\mathbf{x} \bmod p^\ell : g(\mathbf{x}) \equiv 0 \bmod p^\ell\}.$$

Let $M = q^m$ and $\Omega_M = \prod_{p^m || M} \Omega_{p^m}$. Then for $q \mid P(Q)$ (which is necessarily squarefree), by the Chinese Remainder Theorem the condition $\mathbf{x} \bmod p^m \in \Omega_{p^m}^C$ for every $p^m \,||\, M$ may be expressed as $\mathbf{x} \bmod M \in \Omega_M^C$ so that

$$\sum_{\lambda \equiv 0 \bmod q} a_\lambda = \sum_{\substack{\mathbf{x} \in (\mathbb{Z}_{\mathrm{prim}}^{(N+1)})^n \\ F(\mathbf{x}^{(j)})=0 \\ \mathbf{x} \bmod M \in \Omega_M^C}} \prod_{j=1}^{n} \prod_{i=0}^{N} W\left(\frac{x_i^{(j)}}{B_j}\right).$$

We may now break the $\mathbf{x}$ sum into residue classes modulo M, giving

$$\sum_{\lambda \equiv 0 \bmod q} a_\lambda = \sum_{\mathbf{a} \in \Omega_M^C} \sum_{\substack{\mathbf{y} \in (\mathbb{Z}_{\mathrm{prim}}^{(N+1)})^n \\ F(\mathbf{a}^{(j)} + M\mathbf{y}^{(j)})=0}} \prod_{j=1}^{n} \prod_{i=0}^{N} W\left(\frac{a_i^{(j)} + My_i^{(j)}}{B_j}\right)$$
$$= \sum_{\mathbf{a} \in \Omega_M^C} \prod_{j=1}^{n} \sum_{\substack{\mathbf{y}^{(j)} \in (\mathbb{Z}_{\mathrm{prim}}^{(N+1)})^n \\ F(\mathbf{a}^{(j)} + M\mathbf{y}^{(j)})=0}} \prod_{i=0}^{N} W\left(\frac{a_i^{(j)} + My_i^{(j)}}{B_j}\right).$$

Now this is in a form where we may apply 5.2.2, setting $R = \frac{B_j}{M}$, $z_i = \frac{a_i^{(j)}}{B_j}$ and $g(\mathbf{y}^{(j)}) = F(\mathbf{a}^{(j)} + M\mathbf{y}^{(j)})$. Therefore the inner sum over $\mathbf{y}^{(j)}$ can be written as

$$\mathfrak{S}\mathfrak{J}_{W,B_j} + O\left(\left(\frac{B_j}{M}\right)^{N-d+\frac{1}{2}} + M^{\frac{5d(N+1)}{2^d(d-1)}-\frac{3d}{2}}\left(\frac{B_j}{M}\right)^{N+1-d+\epsilon-\frac{(N+1)2^{-d}-(d-1)}{4d}}\right).$$

Observe that

$$\mathfrak{J}_{W,B_j} = \frac{B_j^{N+1-d}}{M^{N+1}}\int_{-\infty}^{\infty}\int_{\mathbb{R}^{N+1}} e(\beta F(\mathbf{u}))\prod_{i=0}^{N} W\left(u_i\right)\mathrm{d}\mathbf{u}\mathrm{d}\beta =: \frac{B_j^{N+1-d}}{M^{N+1}}\mathfrak{J}.$$

The local factors σ_p in the singular series are given by

$$\sigma_p((F(\mathbf{a}^{(j)} + p^m\mathbf{y}^{(j)})) = \lim_{\ell\to\infty} p^{-\ell N}\#\{\mathbf{y} \bmod p^\ell : F(\mathbf{a}^{(j)} + p^m\mathbf{y}^{(j)}) \equiv 0 \bmod p^\ell\}.$$

If $p \nmid M$ then $\sigma_p(F(\mathbf{a}^{(j)}+p^m\mathbf{y}^{(j)})) = \sigma_p$, where σ_p is the usual Hardy–Littlewood density associated to F. If p divides M, it cannot divide $\mathrm{disc}(F)$. Let

$$\mathbf{N}(\ell) := \#\{\mathbf{y} \bmod p^\ell : F(\mathbf{a}^{(j)} + p^m\mathbf{y}) \equiv 0 \bmod p^\ell\},$$

it follows via Hensel's lemma that for $\ell > m$ we have $\mathbf{N}(\ell) = p^N\mathbf{N}(\ell-1)$, and thus $\sigma_p(F(\mathbf{a}^{(j)} + p^m\mathbf{y})) = p^m$. Hence, the singular series factorises as

$$\mathfrak{S} = \prod_{p\nmid M}\sigma_p \prod_{p^m||M} p^m,$$

for any $\mathbf{a}^{(j)}$. Therefore

$$\frac{\mathfrak{S}}{M^{N+1}} = \prod_{p\nmid M}\sigma_p \prod_{p^m||M}\frac{1}{p^{mN}}.$$

Taking the product over all j we get a main term of size

$$\left(\prod_{p\nmid M}\sigma_p^n\right)\mathfrak{J}^n \prod_{p^m||M}\frac{\#\Omega_{p^m}^C}{p^{mnN}}(B_1\dots B_n)^{N+1-d}.$$

Therefore, there exists a constant c (depending at most on W and X) such that

$$\sum_{\lambda\equiv 0 \bmod q} a_\lambda = cg(q)(B_1\dots B_n)^{N+1-d} + O\left(r_q(\mathcal{A})\right),$$

where

$$g(q) = \prod_{p|q}\frac{\#\Omega_{p^m}^C}{\sigma_p^n p^{mnN}} = \prod_{p|q}\frac{\#\widehat{X^n}(\mathbb{Z}/p^m\mathbb{Z}) - \#\Omega_{p^m}}{\#\widehat{X^n}(\mathbb{Z}/p^m\mathbb{Z})}.$$

(In the last equality above we have made use of the fact that $\#\widehat{X}^n(\mathbb{Z}/p^m\mathbb{Z}) = \sigma_p^n p^{mnN}$ which is an immediate consequence of Hensel's lemma see e.g. [19, Lemma 2.1] and the fact that $p \nmid \mathrm{Disc}(F)$.) The remainder term $r_q(\mathcal{A})$ is given by $\#\Omega_M^C \left(\frac{B_1 \dots B_n}{M^n}\right)^{N+1-d}$ multiplied by

$$\left(\frac{\min\limits_j\{B_j\}}{M}\right)^{-\frac{1}{2}} M^{-d(n-1)} + M^{\frac{5d(N+1)}{2^d(d-1)} - \frac{3d}{2} - d(n-1)} \left(\frac{\min\limits_j\{B_j\}}{M}\right)^{-\frac{(N+1)2^{-d}-(d-1)}{4d}+\epsilon}.$$

We estimate $\#\Omega_M^C$ using the following simple bound

$$\#\Omega_M^C \ll \prod_{p^m || M} \#\widehat{X^n}(\mathbb{Z}/p^m\mathbb{Z}) \ll M^{nN}.$$

It just remains to compute the error terms

$$\begin{aligned}\frac{(B_1 \dots B_n)^{N+1-d}}{\min\limits_j\{B_j\}^{1/2}} \sum_{q \leqslant Q^2} \tau_3(M) q^{m(d+\frac{1}{2}-n(N+1))} \#\Omega_{q^m}^C \\ \ll_\epsilon \frac{(B_1 \dots B_n)^{N+1-d}}{\min\limits_j\{B_j\}^{1/2}} \sum_{q \leqslant Q^2} q^{m(d+\frac{1}{2}-n)+\frac{\epsilon}{2}} \\ \ll_\epsilon \frac{(B_1 \dots B_n)^{N+1-d}}{\min\limits_j\{B_j\}^{1/2}} Q^{m(2d-2n+1)+2+\epsilon}\end{aligned}$$

and

$$\begin{aligned}&\frac{(B_1 \dots B_n)^{N+1-d}}{\min\limits_j\{B_j\}^{\frac{(N+1)2^{-d}-(d-1)}{4d}-\epsilon}} \sum_{q \leqslant Q^2} \tau_3(M) q^{m\left(\frac{5d(N+1)}{2^d(d-1)} - \frac{d}{2} - n(N+1) + \frac{(N+1)2^{-d}-(d-1)}{4d} - \epsilon\right)} \#\Omega_{q^m}^C \\ &\ll_\epsilon \frac{(B_1 \dots B_n)^{N+1-d}}{\min\limits_j\{B_j\}^{\frac{(N+1)2^{-d}-(d-1)}{4d}-\epsilon}} Q^{2-m(d+2n)+\frac{5md(N+1)}{2^{d-1}(d-1)}+\frac{(N+1)m2^{-d}-m(d-1)}{2d}-\epsilon}.\end{aligned}$$

This finishes the proof of Theorem 5.1.2.

5.3 Applications

In the remainder of this chapter, we will apply Theorem 5.1.2 and Lemma 5.1.1 to various problems arising in Diophantine geometry.

5.3.1 Thin sets

One of the most important problems in Diophantine geometry is Manin's conjecture concerning the distribution of rational points on Fano varieties. Let X be a smooth projective geometrically integral variety over a number field k.

When the anticanonical bundle $-\omega_X$ is ample, it defines an embedding into a projective space and we expect a large number of points of bounded height. Indeed Manin predicted that the number of points of height $\leqslant B$ should be approximately B times a power of $\log B$. Of course, this cannot be true if we count all the rational points on a variety, for example if X is a cubic surface with a line then the number of points on the line is $\gg B^2$. Therefore it is necessary to avoid so-called *accumulating subvarieties*, of which the lines on a cubic surface are a prototypical example. The conjecture therefore predicted that there exists a Zariski open subset $U \subset X$ such that

$$\#\{x \in U(\mathbb{Q}) : H(x) \leqslant B\} \sim c_{X,H} B(\log B)^{\mathrm{rk}(\mathrm{Pic}X)-1}.$$

The leading constant was conjecturally described by Peyre [89] in terms of the geometry of X. However this conjecture as stated above is still not correct. Counter-examples have been discovered by Batyrev–Tschinkel [4], Le Rudulier [74] and Browning–Heath-Brown [17]. In these counter-examples it is demonstrated that the accumulating subvarieties, which one would want to remove, form a Zariski dense set and therefore it is impossible to find an open set U in which Manin's predicted asymptotic holds. To get around this problem it has recently been suggested that instead of removing just a Zariski closed subset, one should remove a thin subset of $X(k)$.

Definition 5.3.1. If there exists a closed subset $Z \subset X$ with $Z \neq X$ and $A \subset Z(k)$ then we refer to A as a thin set of type I. If there is some irreducible variety Y with $\dim(X) = \dim(Y)$ and a generically surjective morphism $\pi : Y \to X$ of degree $\geqslant 2$ with $A \subset \pi(Y(k))$ then A is referred to as a thin set of type II. A thin subset $A \subset X(k)$ is defined to be a finite union of thin sets of type I and type II.

Another motivation for studying thin sets comes from inverse Galois theory.

Definition 5.3.2. A variety V/k is said to satisfy the *Hilbert property* if the set of rational points $V(k)$ is not thin.
A field k is *Hilbertian* if there exists an irreducible variety of dimension $\geqslant 1$ which has the Hilbert property.

Since closed sets are thin the Hilbert property is a birational property. In fact, it can be shown that k is Hilbertian if and only if $\mathbb{P}^1_k$ satisfies the Hilbert property. Thus $\mathbb{R}$ is not a Hilbertian field, indeed consider the degree 2 maps

$$\begin{aligned} \pi_{i,j} : \mathbb{P}^1 &\to \mathbb{P}^1 \\ (x:y) &\mapsto \left((-1)^i x^2 : (-1)^j y^2\right), \end{aligned}$$

for $(i,j) \in \{0,1\}^2$. The set of real points on $\mathbb{P}^1$ is contained within the union of the images of these maps and hence is thin. Similarly $\mathbb{Q}_p$ is not Hilbertian. However any number field is since the points of bounded height in any thin set can be counted and shown to smaller than the total number of rational points of bounded height (this is [109, Theorem 3.4.4] and the ideas of the proof are very

similar to our proof of Theorem 5.3.4). It follows from Conjecure 1.2.20 that the rational points on any smooth unirational variety over a number field are Zariski dense. This would imply that smooth unirational varieties satisfy the Hilbert property. The reason this is important to Galois theory is the following construction.

Lemma 5.3.3 ([109, Proposition 3.3.1])**.** *If V/k satisfies the Hilbert property and there exists a Galois covering $W \to V$ with Galois group G and such that $V = W/G$ and G acts faithfully on W. Then there exists a field extension of k with Galois group G.*

For any finite group G there exists an n such that G acts faithfully on $W = \mathbb{A}^n_k$. The variety $V = W/G$ is k-unirational. Since Conjecture 1.2.20 implies that V has the Hilbert property, it would follow that one can find a field extension of k with Galois group G, resolving the inverse Galois problem.

Our first application of Theorem 5.1.2 is to demonstrate that points in any given thin subset of a hypersurface of large dimension (or a product of several copies of such) are quantitatively sparse.

Theorem 5.3.4. *Let X be the projective hypersurface defined by a smooth homogeneous form $F \in \mathbb{Z}[x_0, \cdots, x_N]$ of degree d and $A \subset X^n(\mathbb{Q})$ a thin set. If $N > 2^d(d-1) - 1$ and H is the anticanonical height function on X^n described at the end of Section 5.1 then $\exists\, \delta, \gamma > 0$ such that*

$$\#\{x \in A : H(x) \leqslant B\} \ll_{A,X} \begin{cases} B^{1-\delta}(\log B)^{\gamma} & \text{if } n = 1, \\ B(\log B)^{n-2+\gamma} & \text{if } n \geqslant 2. \end{cases}$$

Proof. For each p denote by $\overline{A}_p$ the reduction modulo p and by $\overline{A}$ the collection of all these reductions. We start by breaking into dyadic intervals

$$\#\{x \in A : H(x) \leqslant B\} \leqslant \sum_{\substack{B_1 \dots B_n \leqslant B^{\frac{1}{N+1-d}} \\ \text{dyadic}}} \mathbf{N}(X^n, \mathbf{B}, \overline{A}).$$

(Throughout this proof, the subscript *"dyadic"* under a sum indicates that the variables of the sum vary over powers of 2.) It suffices to prove the estimate in Theorem 5.3.4 when A is either a type I or type II thin set. This will follow from the $m = 1$ case of Theorem 5.1.2. If A is a type I thin set then there is some proper, Zariski closed subset $Z \subset X^n$ in which it is contained. By [19, Lemma 3.8] for all primes p, we have $\#Z(\mathbb{F}_p) \ll_Z p^{n(N+1)-1}$. It follows that there exists a constant $c > 0$ such that $\omega_p \geqslant 1 - \frac{c}{p}$, and thus $\frac{\omega_p}{1-\omega_p} \geqslant \frac{p}{c} - 1$. This means

$$G(Q) \geqslant \sum_{q \leqslant Q} \mu^2(q) q \prod_{p \mid q} \left(\frac{1}{c} - \frac{1}{p} \right) \gg_{\epsilon, Z} Q^{2-\epsilon}.$$

Similarly, if A is a type II thin set then [19, lemma 3.8] tells us that there is a finite Galois extension $k/\mathbb{Q}$ and a constant $c' \in (0,1)$ such that for all primes p which split completely in k we have

$$\#\overline{A}_p \leqslant c' p^{n(N+1)} + O\left(p^{n(N+1)-\frac{1}{2}} \right).$$

Hence by Chebotarev's density theorem this implies that there is a positive density, say δ, set of primes $\mathcal{P}$ and a constant $0 < \eta < \frac{1-c'}{c'}$ such that $\frac{\omega_p}{1-\omega_p} \geqslant \eta$, for large enough $p \in \mathcal{P}$. Hence

$$G(Q) \geqslant \sum_{\substack{q \leqslant Q \\ p|q \implies p \in \mathcal{P} \text{ and } p \text{ suff. large}}} \mu^2(q)\eta^{\omega(q)} \gg_{\epsilon,\mathcal{P}} Q(\log Q)^{\eta\delta - 1}.$$

In either case, 5.1.2 (with $m = 1$) implies that

$$\mathbf{N}(X^n, \mathbf{B}, \overline{A}) \ll_{A,X} (B_1 \dots B_n)^{N+1-d} \left(Q^{-1}(\log Q)^{\gamma} + E_1 + E_2\right),$$

where $\gamma = \eta\delta - 1 \in (0,1)$. Now setting $Q = \min_j \{B_j\}^{\delta}$ for $\delta > 0$ sufficiently small gives the bound

$$\#\{x \in A : H(x) \leqslant B\} \ll_{A,X} (\log B)^{\gamma} \sum_{\substack{B_1 \dots B_n \leqslant B^{\frac{1}{N+1-d}} \\ \text{dyadic}}} (B_1 \dots B_n)^{N+1-d} \min_j B_j^{-\delta},$$

from which the result follows. □

This result extends a theorem of Cohen [27, Theorem 2.5] and a theorem of Browning–Loughran [19, Theorem 1.8] for the cases X is projective space, and X a quadric of dimension at least 3, respectively.

The large sieve has frequently been used for applications in probabilistic Galois theory and we give a further application.

Corollary 5.3.5. *Fix $d \geqslant 3$ and $n > 2^d(d-1)$. Then there exists $\delta > 0$ such that*

$$\#\left\{ f_{\mathbf{a}}(x) = a_n x^n + \dots + a_0 \in \mathbb{Z}[x] : \begin{array}{l} |a_i| \leqslant B, \mathrm{Gal}(f) \neq S_n \\ 2a_n^d = (-a_{n-1})^d + \dots + (-a_0)^d \end{array} \right\} \ll B^{n+1-d-\delta}.$$

Here we have not used the anticanonical height function for purely aesthetic reasons and this doesn't affect the proof.

Proof. Let F be the form defining the polynomial condition on the coefficients and X the associated hypersurface. We define a subvariety W of $X \times \mathbb{A}^1$ by

$$(\mathbf{a}, x) \in W \iff f_{\mathbf{a}}(x) = 0.$$

Since no closed fibre of the map $W \to X$ can have a Galois group larger than the Galois group of the generic fibre, we will be able to deduce that the Galois group in the family is generically S_n if it is S_n for a single fibre. This is true because $x^n - x - 1$ is an example of such a fibre by [109, Remark 2, Section 4.4]. Therefore by Hilbert's Irreducibility Theorem [109, Proposition 3.3.5], it follows that the Galois group of each of the polynomials of the form $f_{\mathbf{a}}$ is S_n outside of a thin set and hence Theorem 5.3.4 implies the result. □

We have chosen this particular family because it is fibred over an appropriate hypersurface and contains a closed fibre which we know to have Galois group S_n. There is nothing particularly special about this example and any other family with these two properties would give a similar result.

In [63], the thin point counting theorem was applied to a problem concerning rational curves on X which we describe briefly below. Two points in the complex projective plane uniquely define a line, and five points a conic. Unfortunately no number of points is guaranteed to uniquely determine a degree 3 curve in $\mathbb{P}^2_{\mathbb{C}}$ however through 8 points in general position there are exactly 12 singular plane cubics. In general, for a positive integer e, let N_e denote the finite number of rational plane curves of degree e through $3e-1$ points. It was shown by Zeuthen [125] in 1873 that $N_4 = 620$ and by Ran [95] and Vainsencher [119] over 100 years later, that $N_5 = 87,304$. Subsequently, Kontsevich and Manin [70] proved a general recursive formula

$$N_e = \sum_{\substack{e_A+e_B=e \\ e_A,e_B>0}} N_{e_A} N_{e_B} e_A^2 e_B \left(e_B \binom{3e-4}{3e_A-2} - e_A \binom{3e-4}{3e_A-1} \right),$$

which allowed N_e to be computed for any e. In [63], myself and Holmes asked the following more arithmetical question: if the $3e-1$ points are all in $\mathbb{P}^2(\mathbb{Q})$, how many of the N_e rational curves are defined over $\mathbb{Q}$? For the case $e = 1$ or 2 the answer is trivially all of them. However in general rarely any of the curves are defined over $\mathbb{Q}$.

Theorem 5.3.6. *Let $e > 2$. Then the set of $(3e-1)$-tuples of points in $\mathbb{P}^2(\mathbb{Q})$ where at least one of the N_e rational curves is defined over $\mathbb{Q}$ forms a thin set.*

One need not, of course, only ask such questions for collections of points in the projective plane. One can replace $\mathbb{P}^2$ by a hypersurface X and the same proof applies as long as one knows that the moduli space of marked rational curves of a given degree on X is irreducible, which is guaranteed for hypersurfaces of large dimension by work of Browning–Vishe [23] and Browning–Sawin [22].

Theorem 5.3.7. *Let $X \subset \mathbb{P}^N$ be a smooth hypersurface of degree d such that $(2d-1)2^{d-1} < N$. Fix integers $e > 0$ and $n \geqslant 0$ such that the expression $(N+1-d)e + (N-4) + n = n(N-1)$ holds. Then there exists a thin set $A \subset X^n(k)$ such that for all $P \in X^n(k) \setminus A$, the corresponding set of rational curves of degree e in X through the points in P contains no curve defined over $\mathbb{Q}$.*

Combining Theorem 5.3.4 with Manin's conjecture for X^n (which follows for example from Birch's work [10] and the compatibility of Manin's conjecture with products e.g. [43, Section 1]) we get the following quantitative statement.

Corollary 5.3.8. *Under the above hypotheses, and assuming that $X(\mathbb{Q})$ is non-empty, the proportion of n-tuples of points in $X(\mathbb{Q})$ for which at least one of the rational curves is defined over $\mathbb{Q}$ is 0%.*

5.3.2 Fibrations

We return to the setting of the Loughran–Smeets conjecture on the number of varieties in a family which have local points everywhere (recall Section 1.3). We may use Lemma 5.1.1 to extend the upper bound of Loughran and Smeets to fibrations over a hypersurface.

Theorem 5.3.9. *Suppose that X is a smooth projective hypersurface defined by the vanishing of the homogeneous degree d form $F \in \mathbb{Z}[x_0, \dots, x_n]$ where $n > 2^d(d-1)$. Supose that Y is a non-singular integral variety over $\mathbb{Q}$ which admits a dominant proper map $\pi : Y \to X$ with geometrically integral generic fibre. Then*

$$\#\{x \in X(\mathbb{Q}) : H(x) \leqslant B \text{ and } \pi^{-1}(x)(\mathbb{A}_{\mathbb{Q}}) \neq \varnothing\} \ll \frac{B}{(\log B)^{\Delta(\pi)}},$$

where $\Delta(\pi)$ is the Loughran–Smeets exponent defined in Section 1.3.1.

Note that the statement given here varies slightly from Theorem B in Chapter 1. This is only because we have chosen to use a different height function, in keeping with the notation of this chapter.

Proof. Denote by $N(X, B, \pi)$ the cardinality which we wish to estimate. Choose $\mathcal{Y}$ a model for Y over $\mathbb{Z}$ along with a map $\pi : \mathcal{Y} \to \mathcal{X}$ which restricts to the original map π on X and Y. Then

$$\begin{aligned} N(X, B, \pi) &\leqslant \#\{x \in X(\mathbb{Q}) : H(x) \leqslant B \text{ and } x \in \pi(Y(\mathbb{Q}_p))\, \forall p\} \\ &\leqslant \#\{x \in X(\mathbb{Q}) : H(x) \leqslant B \text{ and } x \bmod p^2 \in \pi(\mathcal{Y}(\mathbb{Z}_p)) \bmod p^2\, \forall p\}. \end{aligned}$$

Letting $m = 2$ and $\Omega_p^2 = \pi(\mathcal{Y}(\mathbb{Z}_p))$, Lemma 5.1.1 gives us

$$N(X, B, \pi) \ll \frac{B}{\min\{G(B^{\delta_1}), B^{\delta_2}\}},$$

for some $\delta_1, \delta_2 > 0$. By [19, Corollary 3.12], $G(Q) \asymp (\log Q)^{\Delta(\pi)}$ and thus setting Q to be a small enough power of B completes the proof. □

Now that the Loughran–Smeets upper bound has been proven, we take the opportunity to formally announce the Loughran–Smeets conjecture in this setting.

Conjecture 5.3.10. *Retain the assumptions of Theorem 5.3.9 and suppose further that there exists an $x \in X(\mathbb{Q})$ such that $\pi^{-1}(x)(\mathbb{A}_{\mathbb{Q}}) \neq \varnothing$ and that the fibre over every codimension 1 point contains at least one irreducible component of multiplicity 1. Then the upper bound in Theorem 5.3.9 is sharp.*

In [114], Sofos–Visse studied a conic bundle fibration over a low degree hypersurface. Let $f_1, f_2 \in \mathbb{Z}[x_0, \dots, x_n]$ be forms of equal and even degree $d > 0$. Assume that the hypersurfaces associated to f_1 and f_2 are both smooth and

irreducible and that the variety defined by the simultaneous vanishing of f_1 and f_2 is a complete intersection. Further, let $\sigma(f_1, f_2)$ denote the dimension of the variety defined by

$$\operatorname{rank}\left(\frac{\partial f_i}{\partial x_j}\right)_{0\leqslant j\leqslant n}^{i=1,2}(\mathbf{x}) \leqslant 1$$

and suppose that $n-\sigma(f_1,f_2) > 3(d-1)2^d - 1$. Then they studied the function $N_{\mathrm{loc}}(X, B, \pi)$ for the fibration

$$\begin{aligned}\mathbb{P}^2 \times \mathbb{P}^n \supset X : \{t_0^2 + t_1^2 = f_1(\mathbf{x})t_2^2\} &\to B : \{f_2(\mathbf{x}) = 0\} \subset \mathbb{P}^n \\ (\mathbf{t};\mathbf{x}) &\mapsto (\mathbf{x}).\end{aligned}$$

Since this is a smooth hypersurface of sufficiently large dimension compared to its degree we may apply Theorem 5.3.9. The only potentially non-split fibre above a divisor lies above $Z = \{f_1(\mathbf{x}) = f_2(\mathbf{x}) = 0\}$. Similarly to the computation in Remark 6.1, we have that $\delta_Z(\pi) = \frac{1}{2}$ and thus

$$N_{\mathrm{loc}}(X, B, \pi) \ll \frac{B}{(\log B)^{\frac{1}{2}}}.$$

In fact, Sofos–Visse demonstrate that this bound is sharp, confirming the Conjecture in this case, by proving an asymptotic formula of the form

$$N_{\mathrm{loc}}(X, B, \pi) = c\frac{B}{(\log B)^{\frac{1}{2}}} + O\left(\frac{B}{(\log B)^{\frac{1}{2}+\frac{1}{5(d-1)2^{d+5}}}}\right).$$

In general, all currently existing asymptotics for the Loughran–Smeets problem concern conic bundles or more generally Brauer–Severi varieties.

Definition 5.3.11. A Brauer–Severi variety V/k is a variety whose base change to an algebraic closure is isomorphic to projective space, i.e. $V \otimes_k \overline{k} \cong \mathbb{P}^n_{\overline{k}}$.

A Brauer–Severi variety corresponds to an element of the Brauer group of the field over which it is defined. Indeed this is simply a generalisation of the association between conics and quaternion algebras discussed in Section 1.1.1. If $b \in \operatorname{Br} X$ is such that $b(x) = 0 \in \operatorname{Br}\mathbb{Q}$ then the associated Brauer–Severi variety V_b has a rational point. This means that in the setting of products of Brauer–Severi varieties over a hypersurface one could reinterpret Theorem 5.3.9 as a result concerning the vanishing of certain Brauer group elements.

Theorem 5.3.12. *Let X be a smooth projective hypersurface as in Theorem 5.3.9. Let $\mathscr{B} \subset \operatorname{Br} X$ a finite subset. Then*

$$\#\{x \in X(\mathbb{Q}) : H(x) \leqslant B \text{ and } b(x) = 0 \in \operatorname{Br}\mathbb{Q}\ \forall b \in \mathscr{B}\} \ll \frac{B}{(\log B)^{\Delta(\mathscr{B})}}.$$

Here

$$\Delta(\mathscr{B}) = \sum_{D\in X^{(1)}} \frac{1}{|\partial_D(\langle\mathscr{B}\rangle)|},$$

where $\langle\mathscr{B}\rangle$ is the subgroup generated by the elements of $\mathscr{B}$ and $\partial_D : \operatorname{Br} X \to H^1(\kappa(D), \mathbb{Q}/\mathbb{Z})$ the residue map at D.

Proof. Let V be the product of Brauer–Severi varieties associated to the subset $\mathscr{B}$ equipped with the map $\pi : V \to X$. We aim to show that $\delta_D(\pi) = \frac{1}{|\partial_D(\langle\mathscr{B}\rangle)|}$ for each codimension 1 point D in X. Let R be the local ring at D. In [76, Lemma 2.3], a flat proper integral *almost smooth* scheme $\mathcal{V} \xrightarrow{\psi} \operatorname{Spec} R$ is constructed whose generic fibre is isomorphic to the generic fibre of $V \xrightarrow{\pi} X$. (Almost smooth means a scheme $Y \to \operatorname{Spec} R$ such that for any étale R-algebra R' each R'-point of Y lies in the smooth locus of Y.) Moreover the algebraic closure of $\kappa(D)$ in the function field of $\psi^{-1}(D)$ is the compositum of the cyclic extensions generated by the $\partial_D(b)$. Call this compositum K. What this means is that the set of multiplicity one irreducible components of $\psi^{-1}(D)$, the set which we denote $I_D(\pi)$, is isomorphic to the scheme $(\operatorname{Spec} K)^n$ for some $n \in \mathbb{N}$. In [78, Lemma 3.6], it is shown that if Z is integral over F and the algebraic closure L of F in $F(Z)$ is Galois then $\delta(Z) = \frac{1}{[L:F]}$. Each of the Brauer–Severi varieties $V_b/\kappa(D)$ splits over a cyclic extension of degree $\partial_D(b)$. Therefore $K/\kappa(D)$ is Galois and has degree $|\partial_D(\langle\mathscr{B}\rangle)|$. Therefore

$$\delta_D(\psi) = \frac{1}{|\partial_D(\langle\mathscr{B}\rangle)|}.$$

However by [78, Lemma 3.11], since the generic fibres of ψ and π are isomorphic and $\mathcal{V}$ is birational to V we have that

$$\delta_D(\pi) = \delta_D(\psi),$$

which establishes the claim. □

5.3.3 Friable points

One of the most important classes of numbers, and one to which sieve theory frequently finds application, are the smooth or friable numbers. These are integers whose largest prime divisor is not too big. In [19], Browning–Loughran introduced a geometric analogue.

Definition 5.3.13. Let X a scheme of finite type over $\mathbb{Z}$ and $Z \subset X$ a closed subscheme. For $y > 0$ we say that an integral point $x \in X(\mathbb{Z})$ is y-friable with respect to Z if all primes p with $x \bmod p \in Z(\mathbb{F}_p)$ satisfy $p \leqslant y$.

In the case that $X = \mathbb{A}^1_{\mathbb{Z}}$ and Z is the origin, we recover the traditional definition of a y-friable number. Setting $X = \mathbb{A}^1_{\mathbb{Z}}$ and choosing $Z = \{f(x) = 0\}$, means counting those values of x for which the polynomial value $f(x)$ is y-friable. This is a very classical area of study in analytic number theory (the introduction to [71] contains an extensive survey).

We establish the following upper bound, showing the paucity of these friable points.

Theorem 5.3.14. *Let $y > 0$, X a hypersurface as in Theorem 5.3.9 and let $Z \subset X$ be a divisor. Then*

$$\#\{x \in X(\mathbb{Q}) : H(x) \leqslant B \text{ and } x \text{ is } y\text{-friable with respect to } Z\} \ll_{X,y,Z} \frac{B}{(\log B)^{r(Z)}},$$

where $r(Z)$ is the number of irreducible components of Z.

As observed in [19] it is strictly necessary that in general one asks about being friable with respect to a divisor as if we allow a higher codimension subscheme it is possible that a positive proportion of the integral points are friable.

Proof. We will apply Lemma 5.1.1 with $m = 1$ and $\Omega_p = X(\mathbb{F}_p) \setminus Z(\mathbb{F}_p)$ for $p > y$. We have

$$\omega_p = \frac{\#Z(\mathbb{F}_p)}{\#X(\mathbb{F}_p)},$$

and

$$G(Q) \geqslant \sum_{\substack{q \leqslant Q \\ p|q \implies p>y}} \mu^2(q) \prod_{p|q} \omega_p.$$

To evaluate $G(Q)$ we apply Wirsing's theorem [123, Satz 1.1] to the multiplicative function $g(q) = \mu^2(q)q\prod_{p|q}\omega_p$. Since, by [19, Equation 3.11] and an application of partial summation, we have

$$\sum_{p \leqslant Q} \frac{g(p)\log p}{p} = \sum_{p \leqslant Q} \omega_p \log p = r(Z)\log Q + O(1),$$

it follows that

$$\sum_{q \leqslant Q} g(q) \asymp \frac{Q}{\log Q} \prod_{p \leqslant Q} (1 + \omega_p).$$

The product is $\asymp (\log Q)^{r(Z)}$ by [19, Lemma 3.13] and an application of partial summation to remove the extra q in the sum yields $G(Q) \gg (\log Q)^{r(Z)}$. Picking Q a small power of B now finishes the proof. □

One nice example of the application of this theorem is that when $Z = \{x_0 \dots x_n = 0\}$ we get a quantitative bound on the number of solutions to a smooth homogeneous form in y-friable numbers. An important example of this type of estimate is when $X = \{x_1 + x_2 = x_3\}$ since the range of y in comparison to B gives information on the smoothness exponent κ_0 defined by Lagarias–Soundararajan [72] as a variant on the *abc*-conjecture. Theorem 5.3.14 and [19, Theorem 1.11] represent a very introductory investigation into the y-friable points on a variety. Since the study of friable numbers is such a deep area it would be interesting to see if one sees similar growth phenomena of the number of y-friable points in different ranges of y and how the choice of divisor Z plays a role.

Chapter 6

Soluble fibres in a conic bundle

The following is based on joint work with Efthymios Sofos.

6.1 Introduction

In this chapter, we return to the Loughran–Smeets problem (c.f. Section 1.3.1) and study the family of diagonal conics

$$\mathbb{P}^2 \times \mathbb{P}^2 \supset X = \left\{ \sum_{i=0}^{2} a_i X_i^2 = 0 \right\} \xrightarrow{\pi} \mathbb{P}^2$$
$$(a_0 : a_1 : a_2; X_0 : X_1 : X_2) \mapsto (a_0 : a_1 : a_2).$$

We choose the naive height on $\mathbb{P}^2$, namely if $x = [\mathbf{x}]$ for $\mathbf{x} \in \mathbb{Z}^3_{\text{prim}}$ then $H(x) = \max_i |x_i|$. Recall from the introductory chapter, the quantity

$$N_{\text{loc}}(X, B, \pi) := \left\{ x \in \mathbb{P}^2(\mathbb{Q}) : H(x) \leqslant B \text{ and } \pi^{-1}(x)(\mathbb{A}_{\mathbb{Q}}) \neq \varnothing \right\}.$$

The upper bound $N_{\text{loc}}(X, B, \pi) \ll \frac{B^3}{(\log B)^{\frac{3}{2}}}$ was established by Serre [108] and in the same paper he predicts that this upper bound is sharp (a fact confirmed by Guo [53] and Hooley [64]). Serre's paper [108] was the inspiration for the first asymptotic formula in the Loughran–Smeets problem [76] and subsequently for the conjecture itself. In this chapter we will improve on the work of Serre, Hooley and Guo to give an asymptotic formula for this foundational problem.

Theorem 6.1.1. *We have*

$$N_{loc}(X, B, \pi) = c\frac{B^3}{(\log B)^{3/2}} + O\left(\frac{B^3 (\log\log B)^{3/2}}{(\log B)^{5/2}}\right),$$

where the constant c is positive and given by

$$c := \frac{7}{\Gamma\left(\frac{1}{2}\right)^3} \prod_p \left(1 - \frac{1}{p}\right)^{-\frac{3}{2}} \sigma_p,$$

where σ_p is the proportion of fibres which have a $\mathbb{Q}_p$ point (c.f. Section 6.6.4).

Remark. This logarithmic exponent, it should be noted, is in agreement with the Loughran–Smeets conjecture as outlined in Section 1.3.1. Indeed the fibre above any point $(a_0 : a_1 : a_2) \in \mathbb{P}^2(\mathbb{Q})$ such that $a_0 a_1 a_2 \neq 0$ is a smooth conic and thus split. Hence we need only consider the fibres above the lines $L_i = \{a_i = 0\}$. Above L_i, the fibre is given by an equation of the form $aX^2 + bY^2 = 0$ and therefore geometrically the fibre is a union of two lines. The absolute Galois group of the residue field associated to the codimension one point L_i acts by either fixing or permuting these lines so the finite group Γ is isomorphic to $\mathbb{Z}/2\mathbb{Z}$. Since only one of the elements of Γ has a fixed point, we see that

$$\delta_{L_i}(\pi) = \frac{1}{2}.$$

This means

$$\Delta(\pi) = \sum_{i=0}^{2} (1 - \delta_{L_i}(\pi)) = \frac{3}{2}.$$

There are a number of other ways of interpreting our main theorem. Perhaps the simplest is that it gives a count for the number of diagonal ternary quadratic forms which are isotropic as their coefficients vary in a box.

Corollary 6.1.2. *Let $N_{soluble}(B)$ denote the cardinality of the set*

$$\{\mathbf{a} \in \mathbb{Z}^3 : |a_i| \leqslant B \text{ and } a_1X^2 + a_2Y^2 + a_3Z^2 = 0 \text{ has non-trivial } \mathbb{Q}\text{-solution}\}.$$

Then

$$N_{soluble}(B) = 2\zeta(3)c\frac{B^3}{(\log B)^{\frac{3}{2}}} + O\left(\frac{B^3(\log\log B)^{3/2}}{(\log B)^{5/2}}\right),$$

with c as above.

Proof. We have

$$\begin{aligned} N_{\text{soluble}}(B) &= \sum_{\ell \leqslant B} \sum_{\substack{|\mathbf{a}| \leqslant B \\ \gcd(a_1,a_2,a_3)=\ell \\ a_1X^2+a_2Y^2+a_3Z^2=0 \text{ is isotropic}}} 1 \\ &= 2\sum_{\ell \leqslant B} N_{\text{loc}}(X, B/\ell, \pi). \end{aligned}$$

In this last line we are using the Hasse–Minkowski theorem to equate global solubility with everywhere local solubility. The result now follows immediately from Theorem 6.1.1. □

Another way to interpret the result is as a statement regarding the vanishing of Brauer group elements (c.f. the discussion before Theorem 5.3.12). The equation defining our conics may be rewritten as

$$-a_0a_2X_0^2 - a_1a_2X_1^2 = X_2^2.$$

Hence by the connection between conics and quaternion algebras discussed in Section 1.1.1, the fibre above the point $(a_0 : a_1 : a_2) \in \mathbb{P}^2(\mathbb{Q})$ has a rational point exactly when the associated quaternion algebra $(-a_0a_2, -a_1a_2)_{\mathbb{Q}}$ splits. This allows a reframing of the theorem.

Corollary 6.1.3. *Let $b \in \mathrm{Br}(\mathbb{P}^2)$ be the quaternion algebra such that $b(\mathbf{a}) = (-a_0a_2, -a_1a_2)_{\mathbb{Q}}$. If we set*

$$N(b, B) := \#\{a \in \mathbb{P}^2(\mathbb{Q}) : H_{naive}(\mathbf{a}) \leqslant B \text{ and } b(\mathbf{a}) = 0 \text{ in } \mathrm{Br}\,\mathbb{Q}\},$$

then we have

$$N(b, B) = c\frac{B^3}{(\log B)^{\delta_b}}\left(1 + O\left(\frac{(\log\log B)^{3/2}}{\log B}\right)\right),$$

where

$$\delta_b = \sum_{i=0}^{2} \frac{1}{|\partial_{L_i}(\langle b\rangle)|},$$

for $\langle b\rangle$ the subgroup of $\mathrm{Br}(\mathbb{P}^2)$ generated by b and $\partial_{L_i} : \mathrm{Br}(\mathbb{P}^2) \to H^1(\kappa(L_i), \mathbb{Q}/\mathbb{Z})$ the residue map along the line L_i.

Remark. The exponent given above in terms of the residues associated to the quaternion algebra is exactly the same as the exponent in the main theorem. This is proven in [78], and a similar argument is given in the case of a hypersurface base in Theorem 5.3.12.

The layout of this chapter is as follows. In Sections 6.2 and 6.3, we introduce changes of variables to simplify the problem to a form where the methods outlined in Section 2.2.2 may be applied. Section 6.4 contains several useful results which will be necessary. The main counting result is completed in Section 6.5. Finally in Section 6.6, the constant which appears in the asymptotic formula in Section 6.5 is shown to be of the form claimed in Theorem 6.1.1.

6.2 Initial reduction steps

We start by abusing our notation for the Hilbert symbol from Chapter 2 and denoting

$$\left(\frac{m_0, m_1, m_2}{\mathbb{Q}}\right) = \begin{cases} 1, & \text{if } m_0X_0^2 + m_1X_1^2 + m_2X_2^2 = 0 \text{ has a } \mathbb{Q}\text{-point,} \\ 0 & \text{otherwise.} \end{cases}$$

Then, by definition

$$N_{\mathrm{loc}}(X,B,\pi)=\frac{1}{2}\sum_{\substack{\mathbf{a}\in\mathbb{Z}^3_{\mathrm{prim}}\\ ||\mathbf{a}||_\infty\leqslant B}}\left(\frac{\mathbf{a}}{\mathbb{Q}}\right).$$

Our first step is to introduce new variables which account for square factors and common factors of the coefficients a_i. We will write

$$a_i=b_i^2c_i,\quad \mu^2(c_i)=1,$$

and

$$m_{ij}=\gcd(c_i,c_j).$$

Lemma 6.2.1. *For all $B>1$ and $\mathbf{z}\in\mathbb{R}^2_{>0}$, we have*

$$N_{loc}(X,B,\pi)-\frac{1}{2}\sum_{\substack{\mathbf{b}\in(\mathbb{N}\cap[1,z_1])^3\\ \gcd(b_1,b_2,b_3)=1}}\sum_{\substack{\mathbf{m}=(m_{12},m_{13},m_{23})\in\mathbb{N}^3\\ \forall i\neq j:|m_{ij}|\leqslant z_2\\ (6.2.3)}}\mu^2(m_{12}m_{13}m_{23})\mathcal{N}_{\mathbf{b},\ell,\mathbf{m}}(B)$$
$$\ll\frac{B^3}{\min\{B,z_1,z_2\}},$$

where the implied constant is absolute,

$$\mathcal{N}_{\mathbf{b},\ell,\mathbf{m}}(B):=\sum_{\substack{\mathbf{n}\in(\mathbb{Z}\setminus\{0\})^3\\ (6.2.1),(6.2.2)}}\mu^2(n_1n_2n_3)\left(\frac{m_{12}m_{13}n_1,m_{12}m_{23}n_2,m_{13}m_{23}n_3}{\mathbb{Q}}\right)$$

and the conditions in the summation are

$$\begin{cases}|n_1|\leqslant B/(b_1^2m_{12}m_{13}),\\ |n_2|\leqslant B/(b_2^2m_{12}m_{23}),\\ |n_3|\leqslant B/(b_2^2m_{13}m_{23}),\end{cases}\tag{6.2.1}$$

$$\begin{cases}\gcd(n_1n_2n_3,m_{12}m_{13}m_{23})=1,\\ \gcd(n_1,b_2,b_3)=\gcd(n_2,b_1,b_3)=\gcd(n_3,b_1,b_2)=1,\end{cases}\tag{6.2.2}$$

$$\Big\{\gcd(m_{12},b_3)=\gcd(m_{13},b_2)=\gcd(m_{23},b_1)=1.\tag{6.2.3}$$

While initially daunting, the upshot of this lemma is that we have arranged the coefficients of our conics to now be described by six variables which are squarefree and pairwise coprime. Moreover three of these variables, the m_{ij}, can be bounded by a sufficiently large power of $\log B$.

Proof. Firstly note that introducing the condition that $a_1a_2a_3\neq 0$ contributes an error of size $O\left(B^2\right)$, with an absolute implied constant. Since the solubility of the equation $a_1X_1^2+a_2X_2^2+a_3X_3^2=0$ is independent of square factors of

the coefficients we may write each a_i as the product of a square and square-free part

$$\sum_{\substack{\mathbf{a}\in(\mathbb{Z}\setminus\{0\})^3_{\text{prim}}\\ ||\mathbf{a}||_\infty\leqslant B}}\left(\frac{\mathbf{a}}{\mathbb{Q}}\right)=\sum_{\substack{\mathbf{b}\in\mathbb{N}^3_{\text{prim}}\\ ||\mathbf{b}||_\infty\leqslant\sqrt{B}}}\sum_{\substack{\mathbf{c}\in(\mathbb{Z}\setminus\{0\})^3_{\text{prim}}\\ |c_i|\leqslant B/b_i^2\\ \gcd(c_1b_1,c_2b_2,c_3b_3)=1}}\mu^2(c_1)\mu^2(c_2)\mu^2(c_3)\left(\frac{\mathbf{c}}{\mathbb{Q}}\right).$$

Now we restrict the range of $\mathbf{b}$. For any $j=1,2,3$ and $z_1\geqslant 1$, the contribution towards the above sum of those $b_j>z_1$ is bounded by

$$\sum_{\substack{\mathbf{b}\in\mathbb{N}^3_{\text{prim}}\\ ||\mathbf{b}||_\infty\leqslant\sqrt{B}\\ b_j>z_1}}\prod_{i=1}^{3}\left(1+2\frac{B}{b_i^2}\right)\ll B^3\sum_{\substack{\mathbf{b}\in\mathbb{N}^3_{\text{prim}}\\ b_j>z_1}}\frac{1}{b_1^2b_2^2b_3^2}\ll\frac{B^3}{z_1}.$$

We now introduce new variables to keep track of common factors between the c_i. Define $m_{ij}=\gcd(c_i,c_j)$, then there exists $\mathbf{n}\in(\mathbb{Z}\setminus\{0\})^3$ such that

$$c_1=m_{12}m_{13}n_1,\ c_2=m_{12}m_{23}n_2,\ c_3=m_{13}m_{23}n_3.$$

The summation conditions listed above follow immediately from this change of variables. All that remains is to estimate the contribution when $m_{ij}>z_2$. Using the notation $\{i,j,k\}=\{1,2,3\}$ and $c_i'=c_i/m_{ij}$, we bound the contribution to the $\mathbf{c}$ sum by

$$\sum_{z_2<m_{ij}\leqslant\min\{B/b_i^2,B/b_j^2\}}\sum_{\substack{|c_i'|\leqslant B/m_{ij}b_i^2\\ |c_j'|\leqslant B/m_{ij}b_j^2\\ |c_k|\leqslant B/b_k^2}}1\ll\frac{B^3}{b_1^2b_2^2b_3^2}\sum_{m_{ij}>z_2}\frac{1}{m_{ij}^2}\ll\frac{B^3}{z_2b_1^2b_2^2b_3^2}.$$

This completes the lemma. □

Next we turn our attention to the Hilbert symbol and provide the following decomposition.

Lemma 6.2.2. *For every $\mathbf{m}\in\mathbb{N}^3$ and $\mathbf{n}\in(\mathbb{Z}\setminus\{0\})^3_{prim}$,*

$$\left(\frac{m_{12}m_{13}n_1,m_{12}m_{23}n_2,m_{13}m_{23}n_3}{\mathbb{Q}}\right)$$

is equal to

$$\left(\frac{-n_1n_3,-n_2n_3}{\mathbb{R}}\right)\prod_{\substack{p\mid n_1n_2n_3m_{12}m_{13}m_{23}\\ p\neq 2}}\left(\frac{-n_1m_{23}n_3m_{12},-n_2m_{13}n_3m_{12}}{\mathbb{Q}_p}\right).$$

Proof. Since $m_{ij} > 0$ and $n_i \neq 0$ the change of variables

$$Y_1 = m_{12}m_{13}X_1, Y_2 = m_{12}m_{23}X_2, Y_3 = m_{13}m_{23}X_3$$

is invertible over $\mathbb{Q}$ (and all its completions) and it transforms the curve

$$m_{12}m_{13}n_1X_1^2 + m_{12}m_{23}n_2X_2^2 + m_{13}m_{23}n_3X_3^2 = 0$$

into

$$n_1m_{23}Y_1^2 + n_2m_{13}Y_2^2 + n_3m_{12}Y_3^2 = 0.$$

We can therefore write

$$\left(\frac{m_{12}m_{13}n_1, m_{12}m_{23}n_2, m_{13}m_{23}n_3}{\mathbb{Q}}\right) = \left(\frac{n_1m_{23}, n_2m_{13}, n_3m_{12}}{\mathbb{Q}}\right)$$

and invoke the Hasse–Minkowski theorem to write this as

$$\left(\frac{-n_1n_3, -n_2n_3}{\mathbb{R}}\right) \prod_{p \text{ prime}} \left(\frac{-n_1m_{23}n_3m_{12}, -n_2m_{13}n_3m_{12}}{\mathbb{Q}_p}\right).$$

By the explicit description of the local Hilbert symbol given in Lemma 2.2.2, it is clear that $\left(\frac{-n_1m_{23}, n_3m_{12}, -n_2m_{13}n_3m_{12}}{\mathbb{Q}_p}\right) = 0$ when $p \nmid 2n_1n_2n_3m_{12}m_{13}m_{23}$. Finally we may omit the prime $p = 2$ thanks to the Hilbert product formula (2.2.1). □

Finally, we will combine the previous two lemmas with the explicit description of the local Hilbert symbols and quadratic reciprocity, to get the final form of the sums with which we will work. For $\mathbf{d} := (d_1, d_2, d_3)$ and $\mathbf{h} = (h_{12}, h_{13}, h_{23}) \in \mathbb{N}^3$ with $2 \nmid d_1d_2d_3h_{12}h_{13}h_{23}$ we define

$$\begin{aligned} G_{\mathbf{h}}(\mathbf{d}) :=& (d_1-1)(d_2-1) + (d_1-1)(d_3-1) + (d_2-1)(d_3-1) \\ &+ (d_1-1)(h_{12}-1) + (d_1-1)(h_{13}-1) \\ &+ (d_2-1)(h_{12}-1) + (d_2-1)(h_{23}-1) \\ &+ (d_3-1)(h_{13}-1) + (d_3-1)(h_{23}-1) \end{aligned} \tag{6.2.4}$$

and

$$G(\mathbf{h}) := (h_{12}-1)(h_{13}-1) + (h_{12}-1)(h_{23}-1) + (h_{13}-1)(h_{23}-1).$$

For $m \in \mathbb{Z} \setminus \{0\}$ let

$$\tau_{\text{odd}}(m) := \tau(m2^{-v_2(m)}).$$

Lemma 6.2.3. *For $B, z_1, z_2 \geqslant 1$ and $\mathbf{b} \in \mathbb{N}^3_{prim}$, we have*

$$\sum_{\substack{\mathbf{m}\in\mathbb{N}^3 \\ \forall i\neq j: |m_{ij}|\leqslant z_2 \\ (6.2.3)}} \mu^2(m_{12}m_{13}m_{23})\mathcal{N}_{\mathbf{b},\ell,\mathbf{m}}(B) = \sum_{\substack{\mathbf{h},\widetilde{\mathbf{h}}\in\mathbb{N}^3 \\ (6.2.5)}} \frac{(-1)^{\frac{G(\mathbf{h})}{4}}\mathcal{M}_{\mathbf{b},\mathbf{h},\widetilde{\mathbf{h}}}(B)}{\tau_{odd}(h_{12}\widetilde{h_{12}}h_{13}\widetilde{h_{13}}h_{23}\widetilde{h_{23}})},$$

where

$$\begin{cases} \gcd(h_{12}\widetilde{h_{12}}, b_3) = \gcd(h_{13}\widetilde{h_{13}}, b_2) = \gcd(h_{23}\widetilde{h_{23}}, b_1) = 1, \\ 2 \nmid h_{12}h_{13}h_{23}, \mu(h_{12}\widetilde{h_{12}}h_{13}\widetilde{h_{13}}h_{23}\widetilde{h_{23}})^2 = 1, \\ \forall i \neq j : h_{ij}\widetilde{h_{ij}} \leqslant z_2 \end{cases} \tag{6.2.5}$$

and $\mathcal{M}_{\mathbf{b},\mathbf{h},\widetilde{\mathbf{h}}}(B)$ *is defined by*

$$\sum_{\substack{(d_1,d_2,d_3)\in\mathbb{N}^3 \\ (\widetilde{d_1},\widetilde{d_2},\widetilde{d_3})\in(\mathbb{Z}\setminus\{0\})^3 \\ (6.2.6),(6.2.7),(6.2.8)}} \frac{(-1)^{\frac{G_{\mathbf{h}}(\mathbf{d})}{4}}}{\tau_{odd}(d_1\widetilde{d_1}d_2\widetilde{d_2}d_3\widetilde{d_3})} \times \left(\frac{-\widetilde{d_2}\widetilde{d_3}\widetilde{h_{12}}\widetilde{h_{13}}}{d_1h_{23}}\right)\left(\frac{-\widetilde{d_1}\widetilde{d_3}\widetilde{h_{12}}\widetilde{h_{23}}}{d_2h_{13}}\right)\left(\frac{-\widetilde{d_1}\widetilde{d_2}\widetilde{h_{23}}\widetilde{h_{13}}}{d_3h_{12}}\right),$$

with

$$\left(\frac{-\widetilde{d_1}\widetilde{d_3}, -\widetilde{d_2}\widetilde{d_3}}{\mathbb{R}}\right) = 1, \tag{6.2.6}$$

$$\begin{cases} d_1|\widetilde{d_1}| \leqslant B/(b_1^2h_{12}\widetilde{h_{12}}h_{13}\widetilde{h_{13}}), \\ d_2|\widetilde{d_2}| \leqslant B/(b_2^2h_{12}\widetilde{h_{12}}h_{23}\widetilde{h_{23}}), \\ d_3|\widetilde{d_3}| \leqslant B/(b_3^2h_{13}\widetilde{h_{13}}h_{23}\widetilde{h_{23}}), \end{cases} \tag{6.2.7}$$

and

$$\begin{cases} \mu^2(d_1\widetilde{d_1}d_2\widetilde{d_2}d_3\widetilde{d_3}) = 1, 2 \nmid d_1d_2d_3, \\ \gcd(d_1\widetilde{d_1}d_2\widetilde{d_2}d_3\widetilde{d_3}, h_{12}\widetilde{h_{12}}h_{13}\widetilde{h_{13}}h_{23}\widetilde{h_{23}}) = 1, \\ \gcd(d_1\widetilde{d_1}, b_2, b_3) = \gcd(d_2\widetilde{d_2}, b_1, b_3) = \gcd(d_3\widetilde{d_3}, b_1, b_2) = 1. \end{cases} \tag{6.2.8}$$

Again, while notationally intimidating this lemma has a very straightforward interpretation. Following the approach laid out in Section 2.2.2, we are expressing the local Hilbert symbols as Legendre symbols and then expanding the product over primes which divide $n_1n_2n_3m_{12}m_{13}m_{23}$ into a divisor sum. We introduce new variables to denote these divisors and their complements then apply quadratic reciprocity to simplify the product of Legendre symbols, which introduces the $(-1)^{\frac{G_{\mathbf{h}}(\mathbf{d})}{4}}$ and $(-1)^{\frac{G(\mathbf{h})}{4}}$ terms. The convoluted looking summation conditions are simply the result of representing the original size constraints and gcd conditions in terms of these new variables. From hereon out, the reader with vertigo might prefer to think of each of the variables $h_{ij}, \widetilde{h_{ij}}$ and b_i as being equal to 1. Of course, in reality we need to keep track of them.

Proof. By (2.2.2) we can write

$$\prod_{\substack{p|n_im_{jk} \\ p\neq 2}} \left(\frac{-n_1m_{23}n_3m_{12}, -n_2m_{13}n_3m_{12}}{\mathbb{Q}_p}\right)$$

as

$$\frac{1}{\tau_{\mathrm{odd}}(n_i m_{jk})} \sum_{\substack{d_i \in \mathbb{N}, 2\nmid d_i \\ d_i \mid n_i}} \sum_{\substack{h_{jk} \in \mathbb{N}, 2\nmid h_{jk} \\ h_{jk} \mid m_{jk}}} \left(\frac{-n_j n_k m_{ij} m_{ik}}{d_i h_{jk}} \right).$$

Note that here it is crucial that the variables m_{ij} and n_k are each squarefree and are pairwise coprime. We may thus define the integers

$$\widetilde{d_i} := n_i/d_i \text{ and } \widetilde{h_{jk}} := m_{jk}/h_{jk},$$

from which we can infer that

$$\begin{aligned} &\tau_{\mathrm{odd}}(n_1 n_2 n_3 m_{12} m_{13} m_{23}) \prod_{\substack{p \mid n_1 n_2 n_3 m_{12} m_{13} m_{23} \\ p \neq 2}} \left(\frac{-n_1 m_{23} n_3 m_{12}, -n_2 m_{13} n_3 m_{12}}{\mathbb{Q}_p} \right) \\ &= \sum_{\substack{\mathbf{d} \in \mathbb{N}^3 \\ \widetilde{\mathbf{d}} \in (\mathbb{Z} \setminus \{0\})^3 \\ 2 \nmid d_1 d_2 d_3 \\ \forall i : d_i \widetilde{d_i} = n_i}} \sum_{\substack{\mathbf{h} \in \mathbb{N}^3 \\ \widetilde{\mathbf{h}} \in \mathbb{N}^3 \\ 2 \nmid h_{12} h_{13} h_{23} \\ \forall j < k : h_{jk} \widetilde{h_{jk}} = m_{jk}}} \left(\frac{-n_2 n_3 m_{12} m_{13}}{d_1 h_{23}} \right) \left(\frac{-n_1 n_3 m_{12} m_{23}}{d_2 h_{13}} \right) \left(\frac{-n_1 n_2 m_{13} m_{23}}{d_3 h_{12}} \right). \end{aligned} \tag{6.2.9}$$

By the multiplicativity of the Jacobi symbol in the upper argument we obtain

$$\begin{aligned} &\left(\frac{-n_2 n_3 m_{12} m_{13}}{d_1 h_{23}} \right) \left(\frac{-n_1 n_3 m_{12} m_{23}}{d_2 h_{13}} \right) \left(\frac{-n_1 n_2 m_{13} m_{23}}{d_3 h_{12}} \right) \\ &= \left(\frac{-d_2 \widetilde{d_2} d_3 \widetilde{d_3} h_{12} \widetilde{h_{12}} h_{13} \widetilde{h_{13}}}{d_1} \right) \left(\frac{-d_1 \widetilde{d_1} d_3 \widetilde{d_3} h_{12} \widetilde{h_{12}} h_{23} \widetilde{h_{23}}}{d_2} \right) \\ &\times \left(\frac{-d_1 \widetilde{d_1} d_2 \widetilde{d_2} h_{13} \widetilde{h_{13}} h_{23} \widetilde{h_{23}}}{d_3} \right) \left(\frac{-d_2 \widetilde{d_2} d_3 \widetilde{d_3} h_{12} \widetilde{h_{12}} h_{13} \widetilde{h_{13}}}{h_{23}} \right) \\ &\times \left(\frac{-d_1 \widetilde{d_1} d_3 \widetilde{d_3} h_{12} \widetilde{h_{12}} h_{23} \widetilde{h_{23}}}{h_{13}} \right) \left(\frac{-d_1 \widetilde{d_1} d_2 \widetilde{d_2} h_{13} \widetilde{h_{13}} h_{23} \widetilde{h_{23}}}{h_{12}} \right). \end{aligned} \tag{6.2.10}$$

For odd positive integers n, m the reciprocity law for Jacobi symbols is

$$\left(\frac{n}{m} \right) \left(\frac{m}{n} \right) = (-1)^{\frac{(n-1)(m-1)}{4}}.$$

Hence, the right side of (6.2.10) takes the following shape,

$$(-1)^{\frac{G(\mathbf{h}) + G_{\mathbf{h}}(\mathbf{d})}{4}} \left(\frac{-\widetilde{d_2} \widetilde{d_3} \widetilde{h_{12}} \widetilde{h_{13}}}{d_1 h_{23}} \right) \left(\frac{-\widetilde{d_1} \widetilde{d_3} \widetilde{h_{12}} \widetilde{h_{23}}}{d_2 h_{13}} \right) \left(\frac{-\widetilde{d_1} \widetilde{d_2} \widetilde{h_{23}} \widetilde{h_{13}}}{d_3 h_{12}} \right). \tag{6.2.11}$$

The result now follows from combining (6.2.9) and (6.2.11) with Lemmas 6.2.1 and 6.2.2.

□

6.3 Bilinear sums in quadratic characters

At this point our sum has been reduced to one similar, although admittedly more complicated, to (2.2.3). As laid out in Section 2.2.2, we aim to bound $\mathcal{M}_{\mathbf{b},\mathbf{h},\widetilde{\mathbf{h}}}(B)$ by treating each Jacobi symbol as a character and applying Lemma 2.2.4. In this section, we restrict the size of the moduli of these characters so that this plan can be carried out in the following section. The upshot of this section will be that the main term contributions to $\mathcal{M}_{\mathbf{b},\mathbf{h},\widetilde{\mathbf{h}}}(B)$ occur when either all of the d_i are small or all the $\widetilde{d_i}$ are small. This should be compared with the similar situation in Chapter 3.

Firstly we bound the contribution if there are some d_i and $\widetilde{d_j}$ which are both small.

Lemma 6.3.1. *For all $\mathbf{b},\mathbf{h},\widetilde{\mathbf{h}}$ and B as in Lemma 6.2.3, all $z_3 \geqslant 1$ and every $0 < \epsilon < 1/2$ the contribution towards $\mathcal{M}_{\mathbf{b},\mathbf{h},\widetilde{\mathbf{h}}}(B)$ of those $\mathbf{d},\widetilde{\mathbf{d}}$ for which there exists $i \neq j$ with $\max\{d_i, |\widetilde{d_j}|\} \leqslant z_3$ is*

$$\ll_\epsilon \frac{z_3^{1/2-\epsilon}}{B^{1/2-\epsilon}} \frac{B^3}{b_1b_2b_3(h_{12}\widetilde{h_{12}}h_{13}\widetilde{h_{13}}h_{23}\widetilde{h_{23}})^{3/2}},$$

where the implied constant depends at most on ϵ.

Proof. We will only deal with the case $i = 1, j = 2$; the rest of the cases can be treated in an identical manner. The contribution from the range $d_1 \leqslant z_3$ and $|\widetilde{d_2}| \leqslant z_3$ has modulus at most

$$\sum_{\substack{d_1,\widetilde{d_2},d_3,\widetilde{d_3}\\ (6.3.2)}} \frac{\tau_{\mathrm{odd}}(d_1)^{-1}\tau_{\mathrm{odd}}(d_3)^{-1}}{\tau_{\mathrm{odd}}(\widetilde{d_2})\tau_{\mathrm{odd}}(\widetilde{d_3})} \left| \sum_{\substack{\widetilde{d_1}\in\mathbb{Z}\setminus\{0\},d_2\in\mathbb{N}\\ |\widetilde{d_1}|\leqslant B/(d_1b_1^2h_{12}\widetilde{h_{12}}h_{13}\widetilde{h_{13}})\\ d_2\leqslant B/(|\widetilde{d_2}|b_2^2h_{12}\widetilde{h_{12}}h_{23}\widetilde{h_{23}})}} f(\widetilde{d_1})g(d_2)\mu^2(\widetilde{d_1})\mu^2(2d_2)\left(\frac{\widetilde{d_1}}{d_2}\right)\right|, \tag{6.3.1}$$

where the integers $d_1, \widetilde{d_2}, d_3, \widetilde{d_3}$ satisfy

$$1 \leqslant d_1 \leqslant z_3, 1 \leqslant |\widetilde{d_2}| \leqslant z_3, 1 \leqslant d_3|\widetilde{d_3}| \leqslant B/(b_3^2h_{13}\widetilde{h_{13}}h_{23}\widetilde{h_{23}}), \tag{6.3.2}$$

the functions f, g are given through

$$f(\widetilde{d_1}) := \mu^2(\widetilde{h_{12}}h_{13}\widetilde{h_{13}}d_1\widetilde{d_1})\frac{\mathbf{1}_{\mathscr{A}_1}(\widetilde{d_1})}{\tau_{\mathrm{odd}}(\widetilde{d_1})}\left(\frac{\widetilde{d_1}}{d_3h_{12}h_{13}}\right),$$

$$g(d_2) := \frac{\mathbf{1}_{\mathscr{A}_2}(d_2)}{\tau_{\mathrm{odd}}(d_2)}\left(\frac{\widetilde{d_3}\widetilde{h_{12}}\widetilde{h_{23}}}{d_2}\right),$$

the set $\mathscr{A}_1$ consists of all $\widetilde{d_1}$ with

$$\left(\frac{-\widetilde{d_1}\widetilde{d_3}, -\widetilde{d_2}\widetilde{d_3}}{\mathbb{R}}\right) = 1 = \gcd(\widetilde{d_1}, d_1d_3\widetilde{d_2}\widetilde{d_3}h_{12}\widetilde{h_{12}}h_{13}\widetilde{h_{13}}h_{23}\widetilde{h_{23}}\gcd(b_2,b_3))$$

and the set $\mathscr{A}_2$ consists of all d_2 with

$$\gcd(d_2, d_1d_3\widetilde{d_2}\widetilde{d_3}h_{12}\widetilde{h_{12}}h_{13}\widetilde{h_{13}}h_{23}\widetilde{h_{23}}\gcd(b_1,b_3)) = 1.$$

We have not added the condition that $\widetilde{d_1}$ and d_2 be coprime because it is implied by the presence of $(\frac{\widetilde{d_1}}{d_2})$ in (6.3.1). Invoking Lemma 2.2.5 we see that the sum over $\widetilde{d_1}, d_2$ in (6.3.1) is

$$\begin{aligned}\ll_\epsilon &\frac{B^{\frac{3}{2}+\epsilon}}{b_1^2b_2h_{12}^{3/2}\widetilde{h_{12}}^{3/2}h_{13}\widetilde{h_{13}}(h_{23}\widetilde{h_{23}})^{1/2}d_1|\widetilde{d_2}|^{\frac{1}{2}+\epsilon}}\\ &+\frac{B^{\frac{3}{2}+\epsilon}}{b_1b_2^2(h_{12}\widetilde{h_{12}})^{3/2}(h_{13}\widetilde{h_{13}})^{\frac{1}{2}}h_{23}\widetilde{h_{23}}d_1^{\frac{1}{2}+\epsilon}|\widetilde{d_2}|}.\end{aligned}$$

The first term makes the following contribution towards (6.3.1),

$$\begin{aligned}&\ll_\epsilon \frac{B^{\frac{3}{2}+\epsilon}}{b_1^2b_2h_{12}^{3/2}\widetilde{h_{12}}^{3/2}h_{13}\widetilde{h_{13}}(h_{23}\widetilde{h_{23}})^{1/2}}\sum_{\substack{d_1,\widetilde{d_2},d_3,\widetilde{d_3}\\ (6.3.2)}}\frac{\tau_{\mathrm{odd}}(d_1)^{-1}\tau_{\mathrm{odd}}(d_3)^{-1}}{\tau_{\mathrm{odd}}(\widetilde{d_2})\tau_{\mathrm{odd}}(\widetilde{d_3})d_1|\widetilde{d_2}|^{\frac{1}{2}+\epsilon}}\\ &\ll \frac{1}{b_1^2b_2b_3^2}\frac{1}{(h_{12}\widetilde{h_{12}})^{3/2}(h_{13}\widetilde{h_{13}})^2(h_{23}\widetilde{h_{23}})^{3/2}}z_3^{\frac{1}{2}-\epsilon}B^{\frac{5}{2}+\epsilon},\end{aligned}$$

where the standard estimates

$$\begin{aligned}\sum_{d\leqslant x}\frac{1}{d^\alpha\tau_{\mathrm{odd}}(d)} &\ll_\alpha \frac{x^{1-\alpha}}{\sqrt{\log x}},\\ \sum_{d\leqslant x}\frac{1}{d\tau_{\mathrm{odd}}(d)} &\ll \sqrt{\log x},\\ \sum_{dd'\leqslant x}\frac{1}{\tau_{\mathrm{odd}}(dd')} &= \sum_{m\leqslant x}\frac{\tau(m)}{\tau_{\mathrm{odd}}(m)} \ll x,\end{aligned} \tag{6.3.3}$$

have been used for $\alpha = 0$ and $1/2+\epsilon$. Similarly, the second term makes a contribution towards (6.3.1) which is

$$\ll_\epsilon \frac{1}{b_1^2b_2b_3^2}\frac{1}{(h_{12}\widetilde{h_{12}})^{3/2}(h_{13}\widetilde{h_{13}})^2(h_{23}\widetilde{h_{23}})^{3/2}}z_3^{\frac{1}{2}-\epsilon}B^{5/2+\epsilon}.$$

These bounds are satisfactory for our lemma. □

The next complementary lemma is to bound the contribution when a pair of the d_i and $\widetilde{d_j}$ are both large.

Lemma 6.3.2. *For all $\mathbf{b}, \mathbf{h}, \widetilde{\mathbf{h}}$ and B as in Lemma 6.2.3, all $z_3 \geqslant 1$ and every $0 < \epsilon < \frac{1}{2}$ the contribution towards $\mathcal{M}_{\mathbf{b},\mathbf{h},\widetilde{\mathbf{h}}}(B)$ of those $\mathbf{d}, \widetilde{\mathbf{d}}$ for which there exists $i \neq j$ with $\min\{d_i, |\widetilde{d_j}|\} > z_3$ is*

$$\ll_\epsilon \frac{B^3}{b_1 b_2 b_3 h_{12}\widetilde{h_{12}} h_{13}\widetilde{h_{13}} h_{23}\widetilde{h_{23}}} \frac{\sqrt{\log B}}{z_3^{1/2-\epsilon}},$$

where the implied constant depends at most on ϵ.

Proof. We only treat the case $(i,j) = (2,1)$, since the other cases can be dealt with in the same way. By assumption we have $d_2 > z_3$, therefore by (6.2.7) one gets

$$|\widetilde{d_2}| \leqslant B/(z_3 b_2^2 h_{12}\widetilde{h_{12}} h_{23}\widetilde{h_{23}}) \leqslant B/z_3.$$

We similarly obtain $d_1 \leqslant B/z_3$. Hence, the range $\min\{d_2, |\widetilde{d_1}|\} > z_3$ contributes

$$\ll \sum_{\substack{d_1,d_3 \in \mathbb{N} \\ \widetilde{d_2},\widetilde{d_3} \in \mathbb{Z}\setminus\{0\}}} \frac{1}{\tau_{\mathrm{odd}}(d_1)\tau_{\mathrm{odd}}(\widetilde{d_2})\tau_{\mathrm{odd}}(d_3)\tau_{\mathrm{odd}}(\widetilde{d_3})}$$
$$\times \left| \sum_{\substack{\widetilde{d_1} \in \mathbb{Z}\setminus\{0\}, d_2 \in \mathbb{N} \\ |\widetilde{d_1}| \leqslant B/(d_1 b_1^2 h_{12}\widetilde{h_{12}} h_{13}\widetilde{h_{13}}) \\ d_2 \leqslant B/(|\widetilde{d_2}| b_2^2 h_{12}\widetilde{h_{12}} h_{23}\widetilde{h_{23}})}} f'(\widetilde{d_1}) g'(d_2) \mu^2(\widetilde{d_1}) \mu^2(2d_2) \left(\frac{\widetilde{d_1}}{d_2}\right) \right|,$$

where f' and g' are appropriate functions that are analogous to f and g in the proof of Lemma 6.3.1 and the variables $d_1, \widetilde{d_2}, d_3, \widetilde{d_3}$ in the outermost sum satisfy

$$d_1 \leqslant B/z_3, |\widetilde{d_2}| \leqslant B/z_3, d_3|\widetilde{d_3}| \leqslant B/(b_3^2 h_{13}\widetilde{h_{13}} h_{23}\widetilde{h_{23}}).$$

The proof is concluded by using Lemma 2.2.5 and (6.3.3) as in the proof of Lemma 6.3.1. □

We combine these two lemmas to show that there are two main term ranges: when all the d_i are small and the $\widetilde{d_j}$ are large, and vice versa.

Lemma 6.3.3. *For all $B \in \mathbb{R} \cap [1,\infty), \mathbf{z} \in (\mathbb{R} \cap [1,\infty))^4$ and $0 < \epsilon < 1/2$ we have*

$$N_{loc}(X,B,\pi) - \frac{1}{2} \sum_{\substack{\mathbf{b} \in (\mathbb{N}\cap[1,z_1])^3 \\ \gcd(b_1,b_2,b_3)=1}} R_{\mathbf{b}}^{(1)}(B,z_2,z_3) - \frac{1}{2} \sum_{\substack{\mathbf{b} \in (\mathbb{N}\cap[1,z_1])^3 \\ \gcd(b_1,b_2,b_3)=1}} R_{\mathbf{b}}^{(2)}(B,z_2,z_3)$$
$$\ll_\epsilon \frac{B^3}{\min\{B,z_1,z_2\}} + B^{\frac{5}{2}+\epsilon} z_3^{\frac{1}{2}-\epsilon} (\log z_1)^3 + \frac{B^3 (\log z_1)^3 (\log z_2)^3}{z_3^{\frac{1}{2}-\epsilon}},$$

where the implied constant depends at most on ϵ and for $\sigma = 1, 2$ we define

$$R_{\mathbf{b}}^{(\sigma)}(B, z_2, z_3) := \sum_{\substack{\mathbf{h}=(h_{12},h_{13},h_{23})\in\mathbb{N}^3 \\ \widetilde{\mathbf{h}}=(\widetilde{h_{12}},\widetilde{h_{13}},\widetilde{h_{23}})\in\mathbb{N}^3 \\ (6.2.5)}} \frac{(-1)^{\frac{G(\mathbf{h})}{4}}}{\tau_{odd}(h_{12}\widetilde{h_{12}}h_{13}\widetilde{h_{13}}h_{23}\widetilde{h_{23}})} \mathcal{M}_{\mathbf{b},\mathbf{h},\widetilde{\mathbf{h}}}^{(\sigma)}(B, z_3)$$

and let $\mathcal{M}_{\mathbf{b},\mathbf{h},\widetilde{\mathbf{h}}}^{(\sigma)}(B, z_3)$ be given by

$$\sum_{\substack{\mathbf{d}\in\mathbb{N}^3 \\ 2\nmid d_1d_2d_3 \\ \widetilde{\mathbf{d}}\in(\mathbb{Z}\setminus\{0\})^3 \\ (6.2.6),(6.2.7) \\ (6.2.8),(6.3.4)}} \frac{(-1)^{\frac{G_{\mathbf{h}}(\mathbf{d})}{4}}}{\tau_{odd}(d_1\widetilde{d_1}d_2\widetilde{d_2}d_3\widetilde{d_3})} \left(\frac{-\widetilde{d_2}\widetilde{d_3}\widetilde{h_{12}}\widetilde{h_{13}}}{d_1h_{23}}\right)\left(\frac{-\widetilde{d_1}\widetilde{d_3}\widetilde{h_{12}}\widetilde{h_{23}}}{d_2h_{13}}\right)\left(\frac{-\widetilde{d_1}\widetilde{d_2}\widetilde{h_{23}}\widetilde{h_{13}}}{d_3h_{12}}\right),$$

with

$$\begin{cases} \max\{d_i : 1 \leqslant i \leqslant 3\} \leqslant z_3 < \min\{|\widetilde{d_i}| : 1 \leqslant i \leqslant 3\}, & \text{if } \sigma = 1, \\ \max\{|\widetilde{d_i}| : 1 \leqslant i \leqslant 3\} \leqslant z_3 < \min\{d_i : 1 \leqslant i \leqslant 3\}, & \text{if } \sigma = 2. \end{cases} \tag{6.3.4}$$

This lemma should be compared with Lemma 6.2.1. The big difference is that we've simply seperated the two main term ranges and replaced $\mathcal{M}_{\mathbf{b},\mathbf{h},\widetilde{\mathbf{h}}}(B)$ by the same quantity with the appropriate range restriction.

Proof. If any of the d_i and $\widetilde{d_j}$ are in the ranges covered by Lemmas 6.3.1 and 6.3.2 then the contribution is covered by the error term in the statement. Thus we'll restrict to the remaining ranges. Now if $d_1 \leqslant z_3$ then by Lemma 6.3.1 one must have $\min\{|\widetilde{d_2}|, |\widetilde{d_3}|\} > z_3$, which, by Lemma 6.3.2 proves that $\max\{d_2, d_3\} \leqslant z_3$ and therefore $|\widetilde{d_1}| > z_3$ holds due to Lemma 6.3.1. Hence, the condition $d_1 \leqslant z_3$ is equivalent to the case $\sigma = 1$ in (6.3.4). A similar reasoning shows that the condition $d_1 > z_3$ is equivalent to the case $\sigma = 2$ in (6.3.4). □

Note that ultimately, we will be setting each of the z_i to be a large power of $\log B$. Therefore each of the error terms above does constitute a negligible contribution.

6.4 Auxilliary lemmata

The purpose of this section is to gather together several variants of Lemma 2.2.4 which will be useful in our estimation of the sums $\mathcal{M}_{\mathbf{b},\mathbf{h},\widetilde{\mathbf{h}}}^{(\sigma)}(B)$ in the sequel.

The first such result is a version of Lemma 2.2.4 where the function $\tau(n)$ is replaced by $\tau_{\text{odd}}(n)$ as appears in Lemma 6.2.3.

Lemma 6.4.1. *For all Dirichlet characters χ modulo q, all $d \in \mathbb{N}$ and all $C > 0$ we have*

$$\sum_{\substack{1 \leqslant n \leqslant x \\ \gcd(n,d)=1}} \frac{\mu^2(n)\chi(n)}{\tau_{odd}(n)} = \delta_\chi c_0 F(2dq) \left\{ \frac{\mathbf{1}(2 \nmid dq)}{2} + 1 \right\} \frac{x}{\sqrt{\log x}} + \delta_\chi O\left(\frac{x(\log\log 3dq)^{3/2}}{(\log x)^{3/2}} \right) + O_C\left(\frac{\tau(2d)qx}{(\log x)^C} \right),$$

where the implied constant depends at most on C.

Proof. Note that for all $d \in \mathbb{N}, \chi \bmod q$ and d, q not necessarily coprime we have

$$\sum_{\substack{n \leqslant x \\ (n,d)=1}} \frac{\mu^2(n)\chi(n)}{\tau_{\mathrm{odd}}(n)} = \sum_{\substack{n \text{ even} \\ \gcd(n,d)=1}} \frac{\mu^2(n)\chi(n)}{\tau_{\mathrm{odd}}(n)} + \sum_{\substack{n \text{ odd} \\ \gcd(n,d)=1}} \frac{\mu^2(n)\chi(n)}{\tau(n)}$$
$$= \mathbf{1}(2 \nmid d)\chi(2) \sum_{\substack{m \leqslant x/2 \\ \gcd(m,2d)=1}} \frac{\mu^2(m)\chi(m)}{\tau(m)} + \sum_{\substack{n \leqslant x \\ \gcd(n,2d)=1}} \frac{\mu^2(n)\chi(n)}{\tau(n)}.$$

We now apply Lemma 2.2.4 to the two outmost right sums. Furthermore, we have $\delta_\chi \chi(2) = \delta_\chi \mathbf{1}(2 \nmid q)$ because if $2 \mid q$ then $\chi(2) = 0$. Lastly, if $2 \nmid q$ then the condition $\delta_\chi = 1$ implies $\chi(2) = 1$. □

It will be useful to evaluate similar types of sums which also include in their summands arithmetic functions of average order one. To prove this we record the following fairly standard estimate.

Lemma 6.4.2. *Let C and α be any two real numbers with $C > 0$ and $\alpha > 1$. Assume that $f : \mathbb{N} \to \mathbb{C}$ is a multiplicative function that satisfies the bound $f(m) = O_\epsilon(m^\epsilon)$ for every $\epsilon > 0$. Then for all $x \geqslant 2$ we have*

$$\sum_{1 \leqslant m \leqslant x/2} \frac{f(m)}{m^\alpha} \frac{1}{(\log x/m)^C} = \frac{(\sum_{m \in \mathbb{N}} f(m) m^{-\alpha})}{(\log x)^C} + O\left(\frac{1}{(\log x)^{C+1}} \right).$$

Proof. First we observe that

$$\sum_{\sqrt{x} < m \leqslant x/2} \frac{f(m)}{m^\alpha} \frac{1}{(\log x/m)^C} \ll_C \sum_{m > \sqrt{x}} \frac{|f(m)|}{m^\alpha} \ll \sum_{m > \sqrt{x}} \frac{m^{(\alpha-1)/2}}{m^\alpha} \ll \frac{1}{x^{(\alpha-1)/4}},$$

which is sufficiently small. Let us now use the estimate

$$\frac{1}{(1-z)^C} = 1 + O_C(|z|), (|z| \leqslant 1/2)$$

for $z = (\log m)/(\log x)$. We obtain

$$\sum_{1\leqslant m\leqslant\sqrt{x}} \frac{f(m)}{m^\alpha}\frac{1}{(\log x/m)^C} = \frac{1}{(\log x)^C}\sum_{1\leqslant m\leqslant\sqrt{x}} \frac{f(m)}{m^\alpha}\frac{1}{\left(1-\frac{\log m}{\log x}\right)^C}$$
$$= \frac{1}{(\log x)^C}\left(\sum_{1\leqslant m\leqslant\sqrt{x}} \frac{f(m)}{m^\alpha}\right)$$
$$+ O\left(\frac{1}{(\log x)^{C+1}}\sum_{1\leqslant m\leqslant\sqrt{x}} \frac{|f(m)|\log m}{m^\alpha}\right).$$

The growth assumption on f shows that the sum in the error term is $\ll 1$ and that the sum in the main term can be completed by introducing an error term that tends to 0 polynomially fast with x. □

Lemma 6.4.3. *Let η be a fixed positive constant and assume that $f : \mathbb{N} \to \mathbb{R}$ is a multiplicative function satisfying $|f(p) - 1| \leqslant \eta/p$ for all primes p. Then for all Dirichlet characters χ modulo q and all $d \in \mathbb{N}$ with $\gcd(d, q) = 1$ we have*

$$\sum_{\substack{n\leqslant x\\ \gcd(n,d)=1}} \frac{\mu^2(n)\chi(n)f(n)}{\tau(n)} - \frac{x}{\sqrt{\log x}}\delta_\chi c_0 F(dq)\left(\prod_{p\nmid dq}\left(1+\frac{(f(p)-1)}{1+2p}\right)\right)$$
$$\ll \frac{\delta_\chi x(\log\log(3dqx))^{3/2}}{(\log x)^{3/2}} + \frac{\tau(d)qx}{(\log x)^C},$$

where the implied constants depends at most on C and η.

Proof. We define g through $f = 1 * g$, where $*$ denotes the Dirichlet convolution. Then g is multiplicative and satisfies $|g(p)| \leqslant \eta/p$. Hence we obtain

$$\sum_{\substack{1\leqslant n\leqslant x\\ \gcd(n,d)=1}} \frac{\mu^2(n)\chi(n)}{\tau(n)}f(n) = \sum_{\substack{n\leqslant x\\ \gcd(n,d)=1}} \frac{\mu^2(n)\chi(n)}{\tau(n)}\sum_{\ell\mid n} g(\ell)$$
$$= \sum_{1\leqslant\ell\leqslant x} g(\ell) \sum_{\substack{1\leqslant n\leqslant x\\ \ell\mid n\\ \gcd(n,d)=1}} \frac{\mu^2(n)\chi(n)}{\tau(n)}.$$

Writing $a := n/\ell$ gives

$$\sum_{\substack{1\leqslant\ell\leqslant x\\ \gcd(\ell,d)=1}} \frac{g(\ell)\mu^2(\ell)\chi(\ell)}{\tau(\ell)} \sum_{\substack{a\leqslant x/\ell\\ \gcd(a,d\ell)=1}} \frac{\mu^2(a)\chi(a)}{\tau(a)}.$$

The range $\ell > x/2$ can be safely ignored since it contributes

$$\ll \sum_{x/2<\ell\leqslant x} \frac{\eta^{\omega(\ell)}}{\ell\tau(\ell)} \ll \sum_{\ell\leqslant x}\frac{1}{\ell^{1/2}} \ll x^{1/2}.$$

For the remaining range we employ Lemma 2.2.4, thus obtaining

$$\sum_{\substack{1\leqslant \ell\leqslant x/2\\ \gcd(\ell,d)=1}} \frac{g(\ell)\mu^2(\ell)\chi(\ell)}{\tau(\ell)}\left(\frac{\delta_\chi c_0 F(d\ell)x}{\ell\sqrt{\log x/\ell}}\left\{1+O\left(\frac{(\log\log 3d\ell q)^{\frac{3}{2}}}{\log(x/\ell)}\right)\right\}\right.$$
$$\left.+O_C\left(\frac{\tau(d\ell)qx}{\ell(\log x/\ell)^C}\right)\right).$$

Using the bound $\mu^2(\ell)|g(\ell)| \leqslant \eta^{\omega(\ell)}/\ell \ll_\eta \ell^{-9/10}$ shows that the outmost error term contributes

$$\ll_C qx \sum_{\substack{1\leqslant \ell\leqslant x\\ \gcd(\ell,d)=1}} \frac{|g(\ell)|\mu^2(\ell)}{\tau(\ell)}\frac{\tau(d\ell)}{\ell(\log x/\ell)^C}$$
$$\leqslant qx\tau(d)\sum_{\ell\leqslant x}\frac{1}{\ell^{9/10}}\frac{1}{\ell(\log x/\ell)^C} \ll \frac{\tau(d)qx}{(\log x)^C},$$

where we used $\tau(ab)\leqslant\tau(a)\tau(b)$ and Lemma 6.4.2. Using $c(a)=O(1)$ we see that the remaining error term is

$$\ll \delta_\chi x\sum_{\ell\leqslant x/2}\frac{|g(\ell)|}{\tau(\ell)}\frac{1}{\ell\sqrt{\log(x/\ell)}}\frac{(\log\log 3d\ell q)^{3/2}}{\log(x/\ell)}$$
$$\leqslant \delta_\chi x(\log\log 3dqx)^{3/2}\sum_{\ell\leqslant x/2}\frac{1}{\ell^{\frac{19}{10}}(\log x/\ell)^{\frac{3}{2}}},$$

which by Lemma 6.4.2 is $\ll \delta_\chi x(\log\log 3dqx)^{3/2}(\log x)^{-3/2}$. Noting that $\delta_\chi=1$ forces $\chi(\ell)$ to be the indicator function of the integers ℓ that are coprime to q and using Lemma 6.4.2 with

$$\alpha=2, f(m):=\frac{mg(m)\mu^2(m)F(m)}{\tau(m)}, C=\frac{1}{2}$$

shows that the main term is

$$x\delta_\chi c_0 F(d)\sum_{\substack{1\leqslant\ell\leqslant x/2\\ \gcd(\ell,dq)=1}}\frac{g(\ell)\mu^2(\ell)F(\ell)}{\tau(\ell)\ell\sqrt{\log x/\ell}}$$
$$=x\delta_\chi c_0 F(dq)\left(\frac{\prod_{p\nmid dq}\left(1+\frac{g(p)F(p)}{2p}\right)}{\sqrt{\log x}}+O\left(\frac{1}{(\log x)^{3/2}}\right)\right).$$

This is sufficient for the lemma. □

The following is the amalgam of the previous lemmas and the primary tool in our analysis of $\mathcal{M}^{(1)}_{\mathbf{b},\mathbf{h},\widetilde{\mathbf{h}}}(B)$.

Lemma 6.4.4. *Let η be a fixed positive constant and assume that $f : \mathbb{N} \to \mathbb{C}$ is a multiplicative function satisfying $|f(p) - 1| \leqslant \eta/p$ for all primes p. Then for all $C > 0$ and $d, c_1, \dots, c_k \in \mathbb{N}$ we have*

$$\sum_{\substack{1 \leqslant n \leqslant x \\ \gcd(n,d)=1}} \frac{\mu^2(n) f(n)}{\tau_{odd}(n)} \left(\frac{2 + \sum_{i=1}^k \mathbf{1}(2 \nmid c_i n)}{2} \right)$$

$$= \left(\frac{\mathbf{1}(2 \nmid d) f(2) + 2 + \sum_{i=1}^k \mathbf{1}(2 \nmid c_i)}{2} \right) \frac{c_0 F(2d) x}{(\log x)^{1/2}} \left(\prod_{p \nmid 2d} \left(1 + \frac{(f(p) - 1)}{1 + 2p} \right) \right)$$

$$+ O_\eta \left(\frac{x (\log\log(3dx))^{3/2}}{(\log x)^{3/2}} \right),$$

where the implied constants depends at most on C and η.

Proof. Define

$$\mathscr{C}_{d,f}(T) := \sum_{\substack{n \leqslant x \\ \gcd(n,2d)=1}} \frac{\mu^2(n) f(n)}{\tau(n)}.$$

We see, by a similar argument as was used in the proof of Lemma 6.4.1, that the sum over n in our lemma is

$$\mathbf{1}(2 \nmid d) f(2) \mathscr{C}_{d,f}(T/2) + \mathscr{C}_{d,f}(T) + \sum_{i=1}^k \frac{\mathbf{1}(2 \nmid c_i)}{2} \mathscr{C}_{d,f}(T).$$

We can now conclude the proof by using Lemma 6.4.3 (and picking some large value for C) to estimate $\mathscr{C}_{d,f}(T)$. □

Finally when we tackle $\mathcal{M}^{(2)}_{\mathbf{b},\mathbf{h},\tilde{\mathbf{h}}}$ we will need a variant which allows a congruence condition and where the Dirichlet character in question is a Jacobi symbol of the form $\left(\frac{c}{\cdot}\right)$ for some c.

Lemma 6.4.5. *Let $f : \mathbb{N} \to \mathbb{R}_{>0}$ be a multiplicative function satisfying*

$$|f(p) - 1| \leqslant \frac{\eta}{p}$$

for some constant $\eta > 0$ and for all primes p. Then for $j \in \{1, -1\}$, for all square-free integers $s \neq 0$ and $d \in \mathbb{N}$ with $\gcd(d, s) = 1$ we have

$$\sum_{\substack{n \leqslant x \\ n \equiv j \bmod 4 \\ \gcd(n,d)=1}} \frac{\mu^2(n) f(n)}{\tau(n)} \left(\frac{s}{n}\right) = \mathbf{1}_{\{|s|=1\}} \left(\frac{s}{j}\right) \frac{c_0 F(2d) x}{2\sqrt{\log x}} \left(\prod_{p \nmid 2d} \left(1 + \frac{(f(p) - 1))}{2p + 1} \right) \right)$$

$$+ O\left(\mathbf{1}_{\{|s|=1\}} \frac{x}{(\log x)^{3/2}} (\log\log 3dx)^2 \right) + O_C \left(\frac{\tau(d)|s|x}{(\log x)^C} \right),$$

where the implied constants depend at most on C and η and χ_4 is the non-principal character modulo 4.

Proof. We use a character sum to detect the congruence

$$\sum_{\substack{n\leqslant x\\ n\equiv j \bmod 4\\ \gcd(n,d)=1}} \frac{\mu^2(n)f(n)}{\tau(n)}\left(\frac{s}{n}\right) = \frac{1}{2}\sum_{\chi \bmod 4}\chi(j)\sum_{\substack{n\leqslant x\\ \gcd(n,d)=1}}\frac{\mu^2(n)f(n)}{\tau(n)}\left(\frac{s}{n}\right)\chi(n).$$

Now we want to apply Lemma 6.4.3 to the inner sum. The main term occurs when $\left(\frac{s}{\cdot}\right)\chi(\cdot)$ is principal. This occurs exactly when $\overline{\chi}$ induces $\left(\frac{s}{\cdot}\right)$. By the primitivity of the Dirichlet character associated to the Kronecker symbol, this means $\overline{\chi} = \left(\frac{s}{\cdot}\right)$, which happens exactly when $s = \pm 1$, depending on whether χ is the principal or non-principal character mod 4. □

6.5 Proof of main theorem

In this section to simplify notation, we will use the shorthand $h = h_{12}h_{13}h_{23}$ and $\widetilde{h} = \widetilde{h_{12}}\widetilde{h_{13}}\widetilde{h_{23}}$.

6.5.1 The first main term range

In this section we estimate $R^{(1)}_{\mathbf{b}}(B, z_2, z_3)$, introduced in Lemma 6.3.3. To do this we will first investigate $\mathscr{M}^{(1)}_{\mathbf{b},\mathbf{h},\widetilde{\mathbf{h}}}(B)$ which is essentially an average of character sums in the variables $\widetilde{d_i}$. To make this clearer we will rearrange the Jacobi symbols in the definition of $R^{(1)}_{\mathbf{b}}(B, z_2, z_3)$. Let $d_i' = \left|\widetilde{d_i}\right|$ and $\sigma_i \in \{\pm 1\}$ such that $\widetilde{d_i} = \sigma_i d_i'$. The condition (6.2.6) on $\widetilde{\mathbf{d}}$ concerning real solubility can hence be re-expressed as $\boldsymbol{\sigma} := (\sigma_1, \sigma_2, \sigma_3) \in \{\pm 1\}^3 \setminus \{(1,1,1), (-1,-1,-1)\}$. By multiplicativity of the Jacobi symbol, whenever $\{i,j,k\} = \{1,2,3\}$ one sees that

$$\left(\frac{-\widetilde{d_j}\widetilde{d_k}\widetilde{h_{ij}}\widetilde{h_{ik}}}{d_i h_{jk}}\right) = \left(\frac{\sigma_k}{d_i h_{jk}}\right)\left(\frac{\sigma_j}{d_i h_{jk}}\right)\left(\frac{d_k'}{d_i h_{jk}}\right)\left(\frac{d_j'}{d_i h_{jk}}\right)\left(\frac{-\widetilde{h_{ij}}\widetilde{h_{ik}}}{d_i}\right)\left(\frac{-\widetilde{h_{ij}}\widetilde{h_{ik}}}{h_{jk}}\right).$$

We are thus led to expressing $\mathcal{M}^{(1)}_{\mathbf{b},\mathbf{h},\widetilde{\mathbf{h}}}(B,z_3)$ as

$$\left(\frac{-\widetilde{h_{23}}\widetilde{h_{13}}}{h_{12}}\right)\left(\frac{-\widetilde{h_{12}}\widetilde{h_{23}}}{h_{13}}\right)\left(\frac{-\widetilde{h_{12}}\widetilde{h_{13}}}{h_{23}}\right)$$
$$\times \sum_{\substack{\mathbf{d}\in\mathbb{N}^3,\mu(2d_1d_2d_3)^2=1\\ \gcd(d_1d_2d_3,h\widetilde{h})=1\\ \forall\{i,j,k\}=\{1,2,3\}:\gcd(d_i\widetilde{d_i},b_j,b_k)=1\\ d_i\leqslant\min\{z_3,B/(b_i^2h_{ij}\widetilde{h_{ij}}h_{ik}\widetilde{h_{ik}})\}}} \frac{(-1)^{\frac{G_{\mathbf{h}}(\mathbf{d})}{4}}}{\tau(d_1d_2d_3)}\left(\frac{-\widetilde{h_{12}}\widetilde{h_{13}}}{d_1}\right)\left(\frac{-\widetilde{h_{12}}\widetilde{h_{23}}}{d_2}\right)\left(\frac{-\widetilde{h_{23}}\widetilde{h_{13}}}{d_3}\right)$$
$$\times \sum_{\substack{\boldsymbol{\sigma}\in\{1,-1\}^3\\ \boldsymbol{\sigma}\notin\{(1,1,1),(-1,-1,-1)\}}} \left(\frac{\sigma_1}{d_2h_{13}d_3h_{12}}\right)\left(\frac{\sigma_3}{d_2h_{13}d_1h_{23}}\right)\left(\frac{\sigma_2}{d_1h_{23}d_3h_{12}}\right)$$
$$\times \sum_{\substack{(d_1',d_2',d_3')\in\mathbb{N}^3\\ \gcd(d_1'd_2'd_3',h\widetilde{h})=1\\ \forall\{i,j,k\}=\{1,2,3\}:\gcd(d_i',b_j,b_k)=1\\ z_3<d_i'\leqslant B/(b_i^2h_{ij}\widetilde{h_{ij}}h_{ik}\widetilde{h_{ik}}d_i)}} \frac{\mu^2(d_1'd_2'd_3')}{\tau_{\mathrm{odd}}(d_1'd_2'd_3')}\left(\frac{d_1'}{d_2h_{13}d_3h_{12}}\right)\left(\frac{d_2'}{d_1h_{23}d_3h_{12}}\right)\left(\frac{d_3'}{d_2h_{13}d_1h_{23}}\right).$$

We are now in position to deploy Lemma 6.4.1. This tells us that the main term contribution will occur when the characters of the form $\left(\frac{\cdot}{n}\right)$ are principal, of course this can only happen if $n=1$.

Lemma 6.5.1. *Fix $C>0$ and $\mathbf{z}\in\mathbb{R}^3_{\geqslant 1}$ with $z_1,z_2,z_3\ll B$, $z_3>(\log B)^{20}$. Then for all $\mathbf{b}\in\mathbb{N}^3_{prim}$ and all $\mathbf{h},\widetilde{\mathbf{h}}$ satisfying (6.2.5), we have*

$$\mathcal{M}^{(1)}_{\mathbf{b},\mathbf{h},\widetilde{\mathbf{h}}}(B,z_3)=6\mathbf{1}(h=1)\sum_{\substack{(d_1',d_2',d_3')\in\mathbb{N}^3\\ (6.5.1)}}\frac{\mu^2(d_1'd_2'd_3')}{\tau_{odd}(d_1'd_2'd_3')}$$
$$+O_C\left(\frac{(z_1z_2z_3)^{10}}{(b_1b_2b_3h\widetilde{h})^{3/2}}\frac{B^3}{(\log B)^C}\right),$$

where

$$\begin{cases}\gcd(d_1'd_2'd_3',\widetilde{h})=1\\ \forall\{i,j,k\}=\{1,2,3\}:\gcd(d_i',b_j,b_k)=1\\ \forall\{i,j,k\}=\{1,2,3\}:d_i'\leqslant B/(b_i^2\widetilde{h_{ij}}\widetilde{h_{ik}})\end{cases}\tag{6.5.1}$$

Proof. It suffices to bound the contribution from the terms which satisfy $d_1d_2d_3h\neq 1$. We will focus on the case that $d_2h_{13}d_1h_{23}\neq 1$, the proof in the remaining cases being almost identical. In this case we write the sum over the

d_i' as

$$\sum_{\substack{(d_1',d_2')\in\mathbb{N}^2\\ \forall\{i,j,k\}=\{1,2,3\}:\gcd(d_i',b_j,b_k)=1\\ \gcd(d_1'd_2',h\widetilde{h})=1\\ \forall\{i,j,k\}=\{1,2,3\}:z_3<d_i'\leqslant B/(b_i^2h_{ij}\widetilde{h_{ij}}h_{ik}\widetilde{h_{ik}}d_i)}} \frac{\mu^2(d_1d_2)}{\tau_{\mathrm{odd}}(d_1'd_2')}\left(\frac{d_1'}{d_2h_{13}d_3h_{12}}\right)\left(\frac{d_2'}{d_1h_{23}d_3h_{12}}\right)$$

$$\times \sum_{\substack{d_3'\in\mathbb{N}\\ \gcd(d_3',d_1'd_2'\gcd(b_1,b_2)h\widetilde{h})=1\\ z_3<d_3'\leqslant B/(b_3^2h_{13}\widetilde{h_{13}}h_{23}\widetilde{h_{23}}d_3)}} \frac{\mu^2(d_3')}{\tau_{\mathrm{odd}}(d_3')}\left(\frac{d_3'}{d_2h_{13}d_1h_{23}}\right).$$

Extending all of the sums to include the terms less than z_3 introduces an error of at most $O\left(z_3B^2\right)$. Adding the term with $1\leqslant d_3'\leqslant z_3$ to the sum over d_3' introduces an error term that is clearly $O(z_3)$. The sum over the whole range can be estimated by using Lemma 6.4.1with

$$x=\frac{B}{b_3^2h_{13}\widetilde{h_{13}}h_{23}\widetilde{h_{23}}d_3}, d=d_1'd_2'\gcd(b_1,b_2)\widetilde{h_{12}}\widetilde{h_{13}}h_{23}\widetilde{h_{23}},$$
$$q=d_2h_{13}d_1h_{23}, \chi(n):=\left(\frac{n}{d_2h_{13}d_1h_{23}}\right).$$

The fact that $d_2h_{13}d_1h_{23}$ is odd and positive shows that χ is a primitive Dirichlet character with modulus q. Furthermore, since we are in the case with $d_2h_{13}d_3h_{12}\neq 1$ we know that χ is not principal. Therefore by Lemma 6.4.1 we may bound the sum over d_3' by

$$\tau(2d_1'd_2'\gcd(b_1,b_2)h\widetilde{h})\frac{Bd_2d_1}{b_3^2\widetilde{h_{13}}\widetilde{h_{23}}d_3}\left(\log\frac{B}{b_3^2h_{13}\widetilde{h_{13}}h_{23}\widetilde{h_{23}}d_3}\right)^{-C},$$

where C is any fixed positive constant.

Using the ranges for the variables and the trivial bound $\tau(n)\leqslant n$, this upper bound is

$$\ll_C \frac{\tau(d_1')\tau(d_2')z_1z_2^3z_3^2B}{d_3(\log B)^C}.$$

Bounding the sum over $\mathbf{d}$ trivially, we infer that the contribution of the terms with $d_2h_{13}d_3h_{12}\neq 1$ towards $R^{(1)}_{\mathbf{b},\mathbf{h},\widetilde{\mathbf{h}}}(B,z_3)$ is

$$\ll_C z_3^2\sum_{d_1',d_2'\in\mathbb{N}\cap[1,B]}\tau(d_1')\tau(d_2')\frac{z_1z_2^3z_3^2B}{(\log B)^C}\ll_C z_3^2(B\log B)^2\frac{z_1z_2^3z_3^2B}{d_3(\log B)^C}.$$

Summing over $\mathbf{d}$ gives the bound

$$B^3z_1z_2^2z_3^{4+\epsilon}(\log B)^{-C},$$

which when summed over $\mathbf{b}, \mathbf{h}$ and $\widetilde{\mathbf{h}}$ gives

$$\ll B^3(z_1z_2z_3)^{4+\epsilon}(\log B)^{-C}.$$

The claimed error term follows from this and the bounds $b_i \leqslant z_1$ and $h_{ij}\widetilde{h_{ij}} \leqslant z_2$. The 6 which appears in the main term is from the sum over $\boldsymbol{\sigma}$. □

We now arrive at the technical heart of this chapter.

Lemma 6.5.2. *Let $\mathbf{z} \in \mathbb{R}^3_{>0}$ such that $z_i \leqslant (\log B)^{30}$ for each i. Let $\widetilde{\mathbf{h}} \in \mathbb{N}^3$ satisfying* (6.2.5)*. Then there exists a function $f(\mathbf{b}, \widetilde{\mathbf{h}})$, to be defined later, such that*

$$\begin{aligned}\sum_{\substack{(d_1',d_2',d_3')\in\mathbb{N}^3\\ (6.5.1)}} \frac{\mu^2(d_1'd_2'd_3')}{\tau_{odd}(d_1'd_2'd_3')} &= \frac{4}{7\Gamma\left(\frac{1}{2}\right)^3}\prod_p\left(1-\frac{1}{p}\right)^{\frac{3}{2}}\left(1+\frac{3}{2p}\right)\frac{B^3}{(\log B)^{\frac{3}{2}}}\\ &\times\left(\frac{2+\mathbf{1}(2\nmid\widetilde{h})\sum_{1\leqslant i<j\leqslant 3}\mathbf{1}(2\nmid\gcd(b_i,b_j))}{2}\right)\frac{f(\mathbf{b},\widetilde{\mathbf{h}})}{(b_1b_2b_3\widetilde{h})^2}\\ &+O\left(\frac{B(\log\log B)^{3/2}}{(b_1b_2b_3\widetilde{h})^2(\log B)^{5/2}}\right)\end{aligned}$$

Proof. For the duration of this proof we will denote the triple sum which we aim to estimate by

$$\Sigma_0(\mathbf{b},\mathbf{h},\widetilde{\mathbf{h}}) := \sum_{\substack{(d_1',d_2',d_3')\in\mathbb{N}^3\\ (6.5.1)}} \frac{\mu^2(d_1'd_2'd_3')}{\tau_{\mathrm{odd}}(d_1'd_2'd_3')}.$$

We start by writing

$$\Sigma_0(\mathbf{b},\mathbf{h},\widetilde{\mathbf{h}}) = \sum_{\substack{(d_1',d_2')\in\mathbb{N}^2,\gcd(d_1'd_2',\widetilde{h})=1\\ \gcd(d_1',b_2,b_3)=\gcd(d_2',b_1,b_3)=1\\ d_1'\leqslant B/(b_1^2\widetilde{h_{12}}\widetilde{h_{13}}),d_2'\leqslant B/(b_2^2\widetilde{h_{12}}\widetilde{h_{23}})}} \frac{\mu^2(d_1'd_2')}{\tau_{\mathrm{odd}}(d_1'd_2')}\Sigma_1(\mathbf{b},\mathbf{h},\widetilde{\mathbf{h}};d_1',d_2'), \tag{6.5.2}$$

where

$$\Sigma_1(\mathbf{b},\mathbf{h},\widetilde{\mathbf{h}};d_1',d_2') := \sum_{\substack{d_3'\in\mathbb{N},\gcd(d_3',b_1,b_2)=1\\ \gcd(d_3',d_1'd_2'\widetilde{h})=1\\ d_3'\leqslant B/(b_3^2\widetilde{h_{13}}\widetilde{h_{23}})}} \frac{\mu^2(d_3')}{\tau_{\mathrm{odd}}(d_3')}.$$

We estimate Σ_1 using Lemma 6.4.1 with $d = \gcd(b_1,b_2)d_1'd_2'\widetilde{h}$, $C = 5/2$ and

$q = 1$. This yields

$$\begin{aligned}\Sigma_1 =& c(2\gcd(b_1,b_2)d_1'd_2'\widetilde{h})\left\{\frac{\mathbf{1}(2\nmid \gcd(b_1,b_2)d_1'd_2'\widetilde{h})}{2}+1\right\}\\ &\times\frac{B/(b_3^2\widetilde{h_{13}}\widetilde{h_{23}})}{\sqrt{\log(B/(b_3^2\widetilde{h_{13}}\widetilde{h_{23}}))}}\\ &+O\left(c_0F(2\gcd(b_1,b_2)d_1'd_2'\widetilde{h})\frac{B(\log\log(d_1'd_2'\widetilde{h}))^{3/2}}{b_3^2\widetilde{h_{13}}\widetilde{h_{23}}(\log B)^{3/2}}\right).\end{aligned}$$

Using the trivial bound $F(m) := \prod_{p|m}\left(1+\frac{1}{2p}\right)^{-1} = O(1)$ and the bound $\sum_{1\leqslant n\leqslant x}\frac{\mu^2(n)}{\tau_{\mathrm{odd}}(n)} \ll \frac{x}{\sqrt{\log x}}$, we can bound the contribution of the error term by

$$\begin{aligned}&\frac{B(\log\log(B^2\widetilde{h}))^{3/2}}{b_3^2\widetilde{h_{13}}\widetilde{h_{23}}(\log B)^{3/2}}\left(\sum_{\substack{1\leqslant d_1'\leqslant B/(b_1^2\widetilde{h_{12}}\widetilde{h_{13}})\\ 1\leqslant d_2'\leqslant B/(b_2^2\widetilde{h_{12}}\widetilde{h_{23}})}}\frac{\mu^2(d_1')}{\tau_{\mathrm{odd}}(d_1')}\frac{\mu^2(d_2')}{\tau_{\mathrm{odd}}(d_2')}\right)\\ \ll&\frac{B^3(\log\log(Bz_2))^{3/2}}{(b_1b_2b_3\widetilde{h})^2(\log B)^{5/2}}.\end{aligned}$$

Finally note that the error incurred by replacing the $(\log(B/b_3^2\widetilde{h_{13}}\widetilde{h_{23}}))^{-1/2}$ in the main term by $(\log B)^{-1/2}$ is of size at most $\frac{\log\log B}{(\log B)^{3/2}}$. Hence when summed over d_1' and d_2' this leads to an error of size

$$\ll \frac{B^3(\log\log B)}{(b_1b_2b_3\widetilde{h})^2(\log B)^{5/2}}.$$

This means that we have

$$\begin{aligned}\Sigma_0 =&\frac{c_0B}{b_3^2\widetilde{h_{13}}\widetilde{h_{23}}(\log B)^{1/2}}\sum_{\substack{(d_1',d_2')\in\mathbb{N}^2\\ \gcd(d_1'd_2',\widetilde{h})=1\\ \gcd(d_1',b_2,b_3)=\gcd(d_2',b_1,b_3)=1\\ d_1'\leqslant B/(b_1^2\widetilde{h_{12}}\widetilde{h_{13}})\\ d_2'\leqslant B/(b_2^2\widetilde{h_{12}}\widetilde{h_{23}})}}\frac{\mu^2(d_1'd_2')F(2\gcd(b_1,b_2)d_1'd_2'\widetilde{h})}{\tau_{\mathrm{odd}}(d_1'd_2')}\\ &\times\left\{\frac{\mathbf{1}(2\nmid\gcd(b_1,b_2)d_1'd_2'\widetilde{h})}{2}+1\right\}+O\left(\frac{B^3(\log\log B)}{(b_1b_2b_3\widetilde{h})^2(\log B)^{5/2}}\right).\end{aligned}$$

By the identity

$$F(ab) = F(a)F\left(\frac{b}{\gcd(a,b)}\right)\quad\text{for } b \text{ squarefree,}\tag{6.5.3}$$

and the coprimality conditions on d_2, we can rewrite the F term as

$$F(2)F_{\mathrm{odd}}(\gcd(b_1,b_2)d_1'\widetilde{h})F_{\mathrm{odd}}\left(\frac{d_2'}{\gcd(d_2',\gcd(b_1,b_2)d_1'\widetilde{h})}\right)$$
$$=\frac{4}{5}F_{\mathrm{odd}}(\gcd(b_1,b_2)d_1'\widetilde{h})F_{\mathrm{odd}}\left(\frac{d_2'}{\gcd(d_2',\gcd(b_1,b_2))}\right).$$

Using this we see that the main term in the expression for Σ_0 above can be written as

$$\frac{4c_0B}{5b_3^2\widetilde{h_{13}}\widetilde{h_{23}}(\log B)^{1/2}}\sum_{\substack{d_1'\leqslant B/(b_1^2\widetilde{h_{12}}\widetilde{h_{13}})\\ \gcd(d_1',b_2,b_3)=1\\ \gcd(d_1',\widetilde{h})=1}}\frac{\mu^2(d_1')}{\tau_{\mathrm{odd}}(d_1')}F_{\mathrm{odd}}(\gcd(b_1,b_2)d_1'\widetilde{h})\Sigma_2(\mathbf{b},\mathbf{h},\widetilde{\mathbf{h}};d_1'),$$

where

$$\Sigma_2:=\sum_{\substack{d_2'\in\mathbb{N}\\ d_2'\leqslant B/(b_2^2\widetilde{h_{12}}\widetilde{h_{23}})\\ \gcd(d_2',d_1'\gcd(b_1,b_3)\widetilde{h})=1}}\frac{\mu^2(d_2')}{\tau_{\mathrm{odd}}(d_2')}F_{\mathrm{odd}}\left(\frac{d_2'}{\gcd(d_2',b_1,b_2)}\right)$$
$$\times\left\{1+\frac{\mathbf{1}(2\nmid\gcd(b_1,b_2)d_1'd_2'\widetilde{h})}{2}\right\}.$$

We estimate the new sum Σ_2 via Lemma 6.4.4 with

$$f(n):=F_{\mathrm{odd}}\left(\frac{n}{\gcd(n,\gcd(b_1,b_2))}\right),c:=\gcd(b_1,b_2)d_1'\widetilde{h}$$
$$\text{and } d:=d_1'\gcd(b_1,b_3)\widetilde{h}.$$

The ensuing error terms can be handled in a manner that is almost identical to the previous step, so we dispense with the details and just record their total contribution to Σ_0, which is

$$\ll\frac{B^2(\log\log B)^{3/2}}{(b_1b_2\widetilde{h_{12}})^{3/2}(\widetilde{h_{13}}\widetilde{h_{23}})^{1/2}(\log B)^{3/2}}.$$

The main term of Σ_2 is given by

$$\frac{4c_0F_{\mathrm{odd}}(d_1'\gcd(b_1,b_3)\widetilde{h})B}{5b_2^2\widetilde{h_{12}}\widetilde{h_{23}}\sqrt{\log B/(b_2^2\widetilde{h_{12}}\widetilde{h_{23}})}}$$
$$\times\prod_{p\nmid 2d_1'\gcd(b_1,b_3)\widetilde{h}}\left(1+\frac{\left(F\left(\frac{p}{\gcd(p,b_1,b_2)}\right)-1\right)}{(1+2p)}\right)$$
$$\times\left(\frac{\mathbf{1}(2\nmid d_1'\gcd(b_1,b_3)\widetilde{h})+\mathbf{1}(2\nmid d_1'\gcd(b_1,b_2)\widetilde{h})+2}{2}\right).$$

Note that we may again replace $(\log B/(b_2^2\widetilde{h_{12}}\widetilde{h_{23}}))^{-1/2}$ by $(\log B)^{-1/2}$ at the cost of an acceptable error term. Now observe that

$$\prod_{p\nmid 2d_1'\gcd(b_1,b_3)\widetilde{h}}\left(1+\frac{\left(F\left(\frac{p}{\gcd(p,b_1,b_2)}\right)-1\right)}{(1+2p)}\right)$$
$$=\prod_p\left(1+\frac{(F(p)-1)}{(1+2p)}\right)\prod_{p\mid 2d_1'\gcd(b_1,b_2)\gcd(b_1,b_3)\widetilde{h}}\left(1+\frac{(F(p)-1)}{1+2p}\right)^{-1}.$$

Since $F(p)=\frac{2p}{2p+1}$, the Euler product

$$\prod_p\left(\frac{1+(F(p)-1)}{(1+2p)}\right)=\prod_p\left(1-\frac{1}{(2p+1)^2}\right),$$

is absolutely convergent and we denote it by $\mathcal{F}_0$. Moreover, if we let

$$P(m)=\prod_{p\mid m}\left(1-\frac{1}{(2p+1)^2}\right)^{-1}$$

we have

$$\prod_{p\mid 2d_1'\gcd(b_1,b_2)\gcd(b_1,b_3)\widetilde{h}}\left(1+\frac{F(p)-1}{1+2p}\right)^{-1}$$
$$=\frac{25}{24}P_{\mathrm{odd}}(d_1'\gcd(b_1,b_2)\gcd(b_1,b_3)\widetilde{h}).$$

Combining this with the identity (6.5.3) and the analogous identity for P, we see that the main term of Σ_0 is

$$\frac{2}{3}\frac{c_0^2\mathcal{F}_0B^2}{b_2^2b_3^2\widetilde{h_{12}}\widetilde{h_{13}}\widetilde{h_{23}}^2(\log B)}F_{\mathrm{odd}}(\gcd(b_1,b_2)\widetilde{h})F_{\mathrm{odd}}(\gcd(b_1,b_3)\widetilde{h})$$
$$P_{\mathrm{odd}}(\gcd(b_1,b_2)\gcd(b_1,b_3)\widetilde{h})\sum_{\substack{z_3<d_1'\leqslant B/(b_1^2\widetilde{h_{12}}\widetilde{h_{13}})\\ \gcd(d_1',\gcd(b_2,b_3)\widetilde{h})=1}}\frac{\mu^2(d_1')}{\tau_{\mathrm{odd}}(d_1')}$$
$$\times F_{\mathrm{odd}}\left(\frac{d_1'}{\gcd(d_1',b_1,b_2)}\right)F_{\mathrm{odd}}\left(\frac{d_1'}{\gcd(d_1',b_1,b_3)}\right)$$
$$\times P_{\mathrm{odd}}\left(\frac{d_1'}{\gcd(d_1',\gcd(b_1,b_2)\gcd(b_1,b_3))}\right)$$
$$\times\left(\frac{\mathbf{1}(2\nmid d_1'\gcd(b_1,b_3)\widetilde{h})+\mathbf{1}(2\nmid d_1'\gcd(b_1,b_2)\widetilde{h})+2}{2}\right).$$

To simplify the notation a little, we'll denote by $H(n)$ the product

$$F_{\mathrm{odd}}\left(\frac{n}{\gcd(n,b_1,b_3)}\right)P_{\mathrm{odd}}\left(\frac{n}{\gcd(n,\gcd(b_1,b_2)\gcd(b_1,b_3))}\right)F_{\mathrm{odd}}\left(\frac{n}{\gcd(n,b_1,b_2)}\right).$$

Now all that remains is to evaluate the d_1' sum. This is accomplished by appealing to Lemma 6.4.4 with

$$k=2, f=H, d=\gcd(b_2,b_3)\widetilde{h},$$
$$c_1=\gcd(b_1,b_3)\widetilde{h}, c_2=\gcd(b_1,b_2)\widetilde{h}.$$

Again we supress the explicit computation of the error terms and instead just record their total contibution which is

$$\ll_C \frac{B(\log\log B)^{3/2}}{b_1^2\widetilde{h_{12}}\widetilde{h_{13}}(\log B)^{3/2}}+\frac{B}{b_1^2\widetilde{h_{12}}\widetilde{h_{13}}(\log B)^{C}}.$$

The main term is

$$\left(\frac{2+\mathbf{1}(2\nmid\widetilde{h})\sum_{1\leqslant i<j\leqslant 3}\mathbf{1}(2\nmid\gcd(b_i,b_j))}{2}\right)$$
$$\frac{c_0F(2)F_{\mathrm{odd}}(\gcd(b_2,b_3)\widetilde{h})B}{b_1^2\widetilde{h_{12}}\widetilde{h_{13}}(\log B)^{1/2}}\prod_{p\nmid 2\gcd(b_2,b_3)\widetilde{h}}\left(1+\frac{(H(p)-1)}{1+2p}\right).$$

Letting

$$R(n)=\prod_{p\mid n}\left(1-\frac{1}{(p+1)(2p+1)}\right)^{-1},$$

we may rewrite the product above as

$$\prod_{p\nmid 2\gcd(b_1,b_2)\gcd(b_1,b_3)\gcd(b_2,b_3)\widetilde{h}}\left(1+\frac{(F^2P(p)-1)}{2p+1}\right)\prod_{\substack{p\nmid 2\gcd(b_2,b_3)\widetilde{h}\\ p\mid\gcd(b_1,b_2)\gcd(b_1,b_3)}}\left(1+\frac{(F(p)-1)}{2p+1}\right)$$
$$=\prod_{p\nmid 2\gcd(b_1,b_2)\gcd(b_1,b_3)\gcd(b_2,b_3)\widetilde{h}}\frac{1}{R(p)}\prod_{\substack{p\nmid 2\gcd(b_2,b_3)\widetilde{h}\\ p\mid\gcd(b_1,b_2)\gcd(b_1,b_3)}}\frac{1}{P(p)}$$
$$=\frac{15}{14}\mathscr{R}_0\frac{R_{\mathrm{odd}}}{P_{\mathrm{odd}}}\left(\frac{\gcd(b_1,b_2)\gcd(b_1,b_3)}{\gcd(\widetilde{h_{12}},b_1,b_2)\gcd(\widetilde{h_{13}},b_1,b_3)}\right)R_{\mathrm{odd}}(\gcd(b_2,b_3)\widetilde{h}),$$

where

$$\mathscr{R}_0=\prod_p\frac{1}{R(p)}.$$

Here we are using the convention that $\frac{R_{\mathrm{odd}}}{P_{\mathrm{odd}}}(n)=\frac{R_{\mathrm{odd}}(n)}{P_{\mathrm{odd}}(n)}$ to compactify the notation. There is a lot of cancellation with the $\frac{R_{\mathrm{odd}}}{P_{\mathrm{odd}}}$ term above and the term $P_{\mathrm{odd}}(\gcd(b_1,b_2)\gcd(b_1,b_3)\widetilde{h})$ arising from the evaluation of Σ_2. Indeed, combining the two we are left with

$$\frac{15}{14}\mathscr{R}_0PR_{\mathrm{odd}}(\widetilde{h})\prod_{1\leqslant i<j\leqslant 3}\frac{R_{\mathrm{odd}}(\gcd(b_i,b_j))}{R_{\mathrm{odd}}(\gcd(\widetilde{h_{ij}},b_i,b_j))}$$

We note the following identity

$$\left(1+\frac{1}{2p}\right)^3\left(1-\frac{1}{(2p+1)^2}\right)\left(1-\frac{1}{(p+1)(2p+1)}\right)=\left(1+\frac{3}{2p}\right)$$

which allows us to describe the local factors of the Euler product $c_0^3\mathscr{F}_0\mathscr{R}_0$. Now we finally have the result claimed, with the function $f(\mathbf{b},\widetilde{\mathbf{h}})$ being given by

$$F^3PR_{\mathrm{odd}}(\widetilde{h_{12}}\widetilde{h_{13}}\widetilde{h_{23}})\prod_{1\leqslant i<j\leqslant 3}\frac{FR_{\mathrm{odd}}(\gcd(b_i,b_j))}{FR_{\mathrm{odd}}(\gcd(\widetilde{h_{ij}},b_i,b_j))}.$$

□

To complete our analysis of $\sum_{\mathbf{b}}R_{\mathbf{b}}^{(1)}(B,z_2,z_3)$ it remains to perform the sums over $\widetilde{\mathbf{h}}$ and $\mathbf{b}$.

Lemma 6.5.3. *For $z_i\leqslant(\log B)^{30}$, we have*

$$\sum_{\substack{\mathbf{b}\in(\mathbb{N}\cap[1,z_1])^3\\ \gcd(b_1,b_2,b_3)=1}}\sum_{\substack{\mathbf{h}\in\mathbb{N}^3\\ \widetilde{\mathbf{h}}\in\mathbb{N}^3\\ (6.2.5)}}\frac{(-1)^{\frac{G(\mathbf{h})}{4}}}{\tau_{odd}(\widetilde{h\mathbf{h}})}\mathcal{M}_{\mathbf{b},\mathbf{h},\widetilde{\mathbf{h}}}^{(1)}(B,z_3)$$

$$=\frac{9}{\Gamma\left(\frac{1}{2}\right)^3}\frac{B^3}{(\log B)^{\frac{3}{2}}}\prod_p\left(1-\frac{1}{p}\right)^{\frac{3}{2}}\frac{(p^2+p+1)(2p^2+p+2)}{2(p^2-1)^2}+O\left(\frac{B^3(\log\log B)^{\frac{3}{2}}}{(\log B)^{\frac{5}{2}}}\right).$$

Proof. Summing the error terms from Lemmas 6.5.1 and 6.5.2, we see that the total error is bounded by

$$\frac{B^3}{(\log B)^{\frac{3}{2}}}\sum_{\mathbf{b}\in(\mathbb{N}\cap[1,z_1])^3}\sum_{\substack{\mathbf{h},\widetilde{\mathbf{h}}\in\mathbb{N}^3\\ h_{ij}\widetilde{h_{ij}}\leqslant z_2}}\frac{1}{(b_1b_2b_3\widetilde{h})^{\frac{3}{2}}\tau(h\widetilde{h})}$$
$$\times\left(\frac{(\log\log B)^{\frac{3}{2}}\mathbf{1}(h=1)}{(b_1b_2b_3\widetilde{h})^{\frac{1}{2}}\log B}+\frac{(z_1z_2z_3)^{10}}{(\log B)^C}\right).$$

The second error term we can bound trivially and take C large enough to see that it's negligible. In the first term the $\widetilde{\mathbf{h}}$ and $\mathbf{b}$ sums are convergent and extending them to infinity introduces an error of size $\frac{1}{z_1^3z_2^3}$, hence this error term is evidently also acceptable. This same argument shows that we can complete the sums in the main term.

We now turn to simplifying the sums which appear in the main term

$$\sum_{\substack{\mathbf{b}\in\mathbb{N}^3\\ \gcd(b_1,b_2,b_3)=1}}\frac{1}{b_1^2b_2^2b_3^2}\sum_{\substack{\widetilde{\mathbf{h}}=(\widetilde{h_{12}},\widetilde{h_{13}},\widetilde{h_{23}})\in\mathbb{N}^3\\ \forall\{i,j,k\}=\{1,2,3\}\ \gcd(\widetilde{h_{ij}},b_k)=1}}\frac{\mu^2(\widetilde{h})}{\tau_{\mathrm{odd}}(\widetilde{h})\widetilde{h}^2}$$
$$\times\left(\frac{2+\mathbf{1}(2\nmid\widetilde{h})\sum_{1\leqslant i<j\leqslant 3}\mathbf{1}(2\nmid\gcd(b_i,b_j))}{2}\right)f(\mathbf{b},\widetilde{\mathbf{h}}).$$

Since this is a convergent sum of multiplicative functions and the gcd constraints are also multiplicative we may write the sum as an infinite product. We start by breaking the sum into its odd part and the 2-adic factor. Let

$$\mathcal{J}_0 = \sum_{\substack{\mathbf{b}\in\mathbb{N}^3\\ \gcd(b_1,b_2,b_3)=1\\ 2\nmid b_1b_2b_3}} \frac{1}{b_1^2b_2^2b_3^2} \sum_{\substack{\widetilde{\mathbf{h}}\in\mathbb{N}^3\\ \forall\{i,j,k\}=\{1,2,3\}\ \gcd(\widetilde{h_{ij}},2b_k)=1}} \frac{\mu^2(\widetilde{h})f(\mathbf{b},\widetilde{\mathbf{h}})}{\tau_{\mathrm{odd}}(\widetilde{h})\widetilde{h}^2}.$$

The 2-adic part is given by

$$\sum_{\substack{\beta_1,\beta_2,\beta_3\in[0,\infty)\\ \min\{\beta_1,\beta_2,\beta_3\}=0}} \frac{1}{4^{\beta_1+\beta_2+\beta_3}} \sum_{\substack{\eta_{12},\eta_{13},\eta_{23}\in\{0,1\}\\ \sum_{i,j}\eta_{ij}=:\eta\in\{0,1\}\\ \min\{\eta_{ij},\beta_k\}=0 \text{ for } \{i,j,k\}=\{1,2,3\}}} \frac{1}{4^\eta} f(\boldsymbol{\beta},\boldsymbol{\eta}),$$

where

$$\begin{aligned} f(\boldsymbol{\beta},\boldsymbol{\eta}) :&= 1+\frac{\mathbf{1}(\eta=0)\sum_{i,j}\mathbf{1}(\min\{\beta_i,\beta_j\}=0)}{2}\\ &= \begin{cases} 1 \text{ if } \eta=1,\\ \frac{5}{2} \text{ if } \beta_i\geqslant 1 \text{ for at most 1 } i \text{ and } \eta=0,\\ 2 \text{ otherwise.}\end{cases}\end{aligned}$$

The 2-adic part then can be computed as

$$\frac{5}{2}+3\times\frac{1}{4}+3\times\sum_{\beta_k\geqslant 1}\frac{1}{4^{\beta_k}}\left(\frac{1}{4}+\frac{1}{4}+\frac{5}{2}\right)+3\sum_{\beta_i,\beta_j\geqslant 1}\frac{1}{4^{\beta_i+\beta_j}}\left(2+\frac{1}{4}\right)=7$$

Now all we need to do is compute $\mathcal{J}_0$. We write $\mathcal{J}_0$ explicitly as

$$\sum_{\substack{\mathbf{b}\in\mathbb{N}^3\\ \gcd(b_1,b_2,b_3)=1\\ 2\nmid b_1b_2b_3}} \frac{FR(\gcd(b_1,b_2)\gcd(b_1,b_3)\gcd(b_2,b_3))}{b_1^2b_2^2b_3^2}\Lambda(\mathbf{b}),$$

where $\Lambda(\mathbf{b})$ is given by

$$\sum_{\substack{\widetilde{\mathbf{h}}\in\mathbb{N}^3\\ \gcd(\widetilde{h_{ij}},2b_k)=1}} \frac{\mu^2(\widetilde{h})}{\tau(\widetilde{h})}\frac{F^3PR(\widetilde{h})}{\widetilde{h}^2}\prod_{i<j}\frac{1}{FR(\gcd\widetilde{h_{ij}},b_i,b_j)}.$$

This function of $\mathbf{b}$ is best understood as an Euler product

$$\Lambda(\mathbf{b})=\prod_{p\neq 2}\Lambda_p(\mathbf{b}),$$

where

$$\Lambda_p(\mathbf{b}) = \begin{cases} 1 + \frac{3F^3PR(p)}{2p^2} & \text{if } p \nmid b_1b_2b_3, \\ 1 + \frac{F^2P(p)}{2p^2} & \text{if } p \mid \gcd(b_i, b_j) \text{ for some } i, j, \\ 1 + \frac{F^3PR(p)}{p^2} & \text{if } p \mid b_i, p \nmid b_jb_k \text{ for } \{i, j, k\} = \{1, 2, 3\}. \end{cases}$$

We expand $\mathcal{J}_0$ as an Euler product, namely

$$\prod_{p\neq 2} \left(\sum_{\substack{\beta_1,\beta_2,\beta_3\in[0,\infty) \\ \min\{\beta_1,\beta_2\beta_3\}=1 \\ \exists i,j \text{ s.t. } \min\{\beta_i,\beta_j\}\geqslant 1}} \frac{FR(p)}{p^{2\beta_i+2\beta_j}} \left(1 + \frac{F^2P(p)}{2p^2}\right) \right.$$

$$\left. + \sum_{\substack{\beta_1,\beta_2\beta_3\in[0,\infty) \\ \exists k \text{ s.t. } \beta_k\geqslant 1 \\ \beta_i=\beta_j=0}} \frac{1}{p^{2\beta_k}} \left(1 + \frac{F^3PR(p)}{p^2}\right) + 1 + \frac{3F^3PR(p)}{2p^2} \right).$$

We record here the identities

$$F^3PR(p) = \left(1 + \frac{3}{2p}\right)^{-1}$$
$$F^2P(p) = \left(1 + \frac{1}{p}\right)^{-1}$$
$$FR(p) = \frac{2p+2}{2p+3}.$$

Using these we conclude

$$\begin{aligned} \mathcal{J}_0 &= \prod_{p>2} \left(3\frac{1}{(p^2-1)^2} \frac{2p+2}{2p+3} \left(1 + \frac{1}{2p^2} \frac{1}{1+1/p}\right) \right. \\ &\qquad \left. + 3\frac{1}{p^2-1} \left(1 + \frac{1}{p^2} \frac{1}{1+3/2p}\right) + 1 + \frac{3}{2p^2} \frac{1}{1+3/2p} \right) \\ &= \prod_{p>2} \left(\frac{(p^2+p+1)(2p^2+1+2p)}{2(1+\frac{3}{2p})(p^2+1)^2(p^2-1)^2} \right) \\ &= \frac{3}{8} \prod_{p} \left(\frac{(p^2+p+1)(2p^2+1+2p)}{2(1+\frac{3}{2p})(p^2+1)^2(p^2-1)^2} \right). \end{aligned}$$

Multiplying this by the 2-adic factor 7, the 6 appearing in Lemma 6.5.1 and $\frac{4}{7\Gamma\left(\frac{1}{2}\right)^3} \prod_p \left(1 - \frac{1}{p}\right)^{3/2} \left(1 + \frac{3}{2p}\right)$, from Lemma 6.5.2, completes the claim. □

6.5.2 The second main term range

We now turn to the estimation of $R^{(2)}(B, z_2, z_3)$. This will follow along similar lines to the previous subsection of course. We start by looking at $\mathcal{M}^{(2)}_{\mathbf{b},\mathbf{h},\tilde{\mathbf{h}}}(B, z_3)$, which we write as

$$\sum_{\substack{|\widetilde{d_i}|\leqslant z_3\\ \gcd(\widetilde{d_1}\widetilde{d_2}\widetilde{d_3},h\widetilde{h})=1\\ (\frac{-\widetilde{d_1}\widetilde{d_3},-\widetilde{d_2}\widetilde{d_3}}{\mathbb{R}})=1\\ \forall\{i,j,k\}=\{1,2,3\}\ \gcd(\widetilde{d_k},b_i,b_j)=1}} \frac{\mu^2(\widetilde{d_1}\widetilde{d_2}\widetilde{d_3})}{\tau_{\mathrm{odd}}(\widetilde{d_1}\widetilde{d_2}\widetilde{d_3})}\left(\frac{-\widetilde{d_2}\widetilde{d_3}\widetilde{h_{12}}\widetilde{h_{13}}}{h_{23}}\right)\left(\frac{-\widetilde{d_1}\widetilde{d_3}\widetilde{h_{12}}\widetilde{h_{23}}}{h_{13}}\right)\left(\frac{-\widetilde{d_1}\widetilde{d_2}\widetilde{h_{23}}\widetilde{h_{13}}}{h_{12}}\right)$$

$$\times \sum_{\substack{1\leqslant d_1\leqslant B/(b_1^2h_{12}\widetilde{h_{12}}h_{13}\widetilde{h_{13}}|\widetilde{d_1}|)\\ \gcd(d_1,2\gcd(b_2,b_3)\widetilde{d_1}h\widetilde{h})=1\\ 1\leqslant d_2\leqslant B/b_1^2h_{12}\widetilde{h_{12}}h_{23}\widetilde{h_{23}}|\widetilde{d_2}|\\ \gcd(d_2,2\gcd(b_1,b_3)d_1\widetilde{d_2}h\widetilde{h})=1}} \frac{\mu^2(d_1d_2)}{\tau(d_1d_2)}\left(\frac{-\widetilde{d_2}\widetilde{d_3}\widetilde{h_{12}}\widetilde{h_{13}}}{d_1}\right)\left(\frac{-\widetilde{d_1}\widetilde{d_3}\widetilde{h_{12}}\widetilde{h_{23}}}{d_2}\right)$$

$$\times \sum_{\substack{1\leqslant d_3\leqslant B/b_3^2h_{13}\widetilde{h_{13}}h_{23}\widetilde{h_{23}}|\widetilde{d_3}|\\ \gcd(d_3,2\gcd(b_1,b_2)d_1d_2\widetilde{d_3}h\widetilde{h})=1}} (-1)^{G_{\mathbf{h}}(\mathbf{d})/4}\frac{\mu^2(d_3)}{\tau(d_3)}\left(\frac{-\widetilde{d_1}\widetilde{d_2}\widetilde{h_{13}}\widetilde{h_{23}}}{d_3}\right).$$

There are two differences between the situation in this range compared to the prequel. The first is that the variable of summation now appears in the lower argument of the Jacobi symbol. This is still a Dirichlet character, but the conductor is either the absolute value of the upper argument or 4 times this, depending on congruence conditions modulo 4. The second difference is the presence of the term $(-1)^{G_{\mathbf{h}}(\mathbf{d})/4}$ in the sum. Referring back to the definition (6.2.4), it is clear that this term only depends on the residue of $\mathbf{d}$ modulo 4, so we can remove it from the sum by breaking into residue classes. Hence we see

that $\mathcal{M}^{(2)}_{\mathbf{b},\mathbf{h},\widetilde{\mathbf{h}}}(B,z_3)$ is equal to

$$\sum_{\substack{|\widetilde{d_i}|\leqslant z_3\\ \gcd(\widetilde{d_1}\widetilde{d_2}\widetilde{d_3},h\widetilde{h})=1\\ \left(\frac{-\widetilde{d_1}\widetilde{d_3},-\widetilde{d_2}\widetilde{d_3}}{\mathbb{R}}\right)=1\\ \forall\{i,j,k\}=\{1,2,3\}\ \gcd(\widetilde{d_k},b_i,b_j)=1}}\frac{\mu^2(\widetilde{d_1}\widetilde{d_2}\widetilde{d_3})}{\tau_{\mathrm{odd}}(\widetilde{d_1}\widetilde{d_2}\widetilde{d_3})}\left(\frac{-\widetilde{d_2}\widetilde{d_3}\widetilde{h_{12}}\widetilde{h_{13}}}{h_{23}}\right)\left(\frac{-\widetilde{d_1}\widetilde{d_3}\widetilde{h_{12}}\widetilde{h_{23}}}{h_{13}}\right)\left(\frac{-\widetilde{d_1}\widetilde{d_2}\widetilde{h_{23}}\widetilde{h_{13}}}{h_{12}}\right)$$
$$\times\sum_{\boldsymbol{\epsilon}\in\{1,3\}^3}(-1)^{G_{\mathbf{h}}(\boldsymbol{\epsilon})/4}$$
$$\times\sum_{\substack{1\leqslant d_1\leqslant B/(b_1^2h_{12}\widetilde{h_{12}}h_{13}\widetilde{h_{13}}|\widetilde{d_1}|)\\ \gcd(d_1,2\gcd(b_2,b_3)\widetilde{d_1}h\widetilde{h})=1\\ 1\leqslant d_2\leqslant B/b_1^2h_{12}\widetilde{h_{12}}h_{23}\widetilde{h_{23}}|\widetilde{d_2}|\\ \gcd(d_2,2\gcd(b_1,b_3)d_1\widetilde{d_2}h\widetilde{h})=1\\ (d_1,d_2)\equiv(\epsilon_1,\epsilon_2)\bmod 4}}\frac{\mu^2(d_1d_2)}{\tau(d_1d_2)}\left(\frac{-\widetilde{d_2}\widetilde{d_3}\widetilde{h_{12}}\widetilde{h_{13}}}{d_1}\right)\left(\frac{-\widetilde{d_1}\widetilde{d_3}\widetilde{h_{12}}\widetilde{h_{23}}}{d_2}\right)$$
$$\times\sum_{\substack{1\leqslant d_3\leqslant B/b_3^2h_{13}\widetilde{h_{13}}h_{23}\widetilde{h_{23}}|\widetilde{d_3}|\\ \gcd(d_3,2\gcd(b_1,b_2)d_1d_2\widetilde{d_3}h\widetilde{h})=1\\ d_3\equiv\epsilon_3\bmod 4}}\frac{\mu^2(d_3)}{\tau(d_3)}\left(\frac{-\widetilde{d_1}\widetilde{d_2}\widetilde{h_{13}}\widetilde{h_{23}}}{d_3}\right).$$

Our first step is the analogue of Lemma 6.5.1.

Lemma 6.5.4. *Fix $C>0$ and $\mathbf{z}\in\mathbb{R}^3_{\geqslant 1}$ with $z_1,z_2,z_3\ll B$ and assume $z_3>(\log B)^{20}$. Then for all $\mathbf{b}\in\mathbb{N}^3_{prim}$ and all $\mathbf{h},\widetilde{\mathbf{h}}$ satisfying (6.2.5), we have that $\mathcal{M}^{(2)}_{\mathbf{b},\mathbf{h},\widetilde{\mathbf{h}}}(B,z_3)$ is equal to*

$$\mathbf{1}(\widetilde{h}=1)\sum_{\substack{\boldsymbol{\epsilon}\in\{1,3\}^3\\ \widetilde{\mathbf{d}}\in\{\pm1\}^3\setminus\{(1,1,1),(-1,-1,-1)\}}}(-1)^{G_{\mathbf{h}}(\boldsymbol{\epsilon})/4}\left(\frac{-\widetilde{d_2}\widetilde{d_3}}{\epsilon_1h_{23}}\right)\left(\frac{-\widetilde{d_1}\widetilde{d_3}}{\epsilon_2h_{13}}\right)\left(\frac{-\widetilde{d_1}\widetilde{d_2}}{\epsilon_3h_{12}}\right)$$
$$\times\sum_{\substack{(d_1,d_2,d_3)\in\mathbb{N}^3\\ (6.5.4)}}\frac{\mu^2(d_1d_2d_3)}{\tau_{odd}(d_1d_2d_3)}+O_C\left(\frac{(z_1z_2z_3)^{10}}{(b_1b_2b_3h\widetilde{h})^{3/2}}\frac{B^3}{(\log B)^C}\right),$$

where

$$\begin{cases}\gcd(d_1d_2d_3,h)=1\\ \forall\{i,j,k\}=\{1,2,3\}:\gcd(d_i,b_j,b_k)=1\\ \forall\{i,j,k\}=\{1,2,3\}:d_i\leqslant B/(b_i^2h_{ij}h_{ik})\\ \mathbf{d}\equiv\boldsymbol{\epsilon}\bmod 4\end{cases}\tag{6.5.4}$$

Proof. The proof in this range is similar, using Lemma 6.4.5 instead of Lemma 6.4.1. However by looking at the error terms in these two results it's clear that the computation of the contribution from the terms where $|\widetilde{d_1}\widetilde{d_2}\widetilde{d_3}|\widetilde{h}\neq 1$ is

identical. The main term occurs when $|\widetilde{d_1}\widetilde{d_2}\widetilde{d_3}|\widetilde{h_{12}}\widetilde{h_{13}}\widetilde{h_{23}} = 1$ but each $\widetilde{d_i}$ could be +1 or -1. Either way the value of the symbols $\left(\frac{-\widetilde{d_i}\widetilde{d_j}}{d_k h_{ij}}\right)$ depends at most on the residue class of d_k mod 4. □

We record again the contribution from the innermost sum.

Lemma 6.5.5. *Let $\mathbf{z} \in \mathbb{R}^3_{>0}$ such that $z_i \leqslant (\log B)^{30}$ for each i. Let $\widetilde{\mathbf{h}} \in \mathbb{N}^3$ satisfying* (6.2.5)*. Then there exists a function $f(\mathbf{b}, \widetilde{\mathbf{h}})$, to be defined later, such that*

$$\sum_{\substack{(d_1,d_2,d_3)\in\mathbb{N}^3 \\ (6.5.4)}} \frac{\mu^2(d_1d_2d_3)}{\tau_{odd}(d_1d_2d_3)} = \frac{1}{14\Gamma\left(\frac{1}{2}\right)^3}\prod_p\left(1-\frac{1}{p}\right)^{\frac{3}{2}}\left(1+\frac{3}{2p}\right)\frac{B^3}{(\log B)^{\frac{3}{2}}}$$
$$\times \frac{f(\mathbf{b},\mathbf{h})}{(b_1b_2b_3h)^2}$$
$$+ O\left(\frac{B(\log\log B)^{3/2}}{(b_1b_2b_3h)^2(\log B)^{5/2}}\right) + O_C\left(\frac{B^3}{(b_1b_2b_3h)^{3/2}(\log B)^C}\right).$$

Proof. We use character sums to pick out the congruence conditions mod 4 so that our sum becomes

$$\frac{1}{8}\sum_{\chi_i \bmod 4}\prod_{i=1}^3\chi_i(\epsilon_i) \sum_{\substack{\mathbf{d}\in\mathbb{N}^3 \\ \gcd(d_1d_2d_3,h)=1 \\ \forall\{i,j,k\}=\{1,2,3\}:\gcd(d_i,b_j,b_k)=1 \\ \forall\{i,j,k\}=\{1,2,3\}:d_i\leqslant B/(b_i^2h_{ij}h_{ik})}} \frac{\mu^2(d_1d_2d_3)}{\tau(d_1d_2d_3)}\prod_{i=1}^3\chi_i(d_i).$$

If any of the characters χ_i are non-principal then we may apply the error term of Lemma 6.4.1 to obtain an error of size

$$\ll_C \frac{B^3}{(\log B)^C}\frac{\tau(b_1b_2b_3h)}{(b_1b_2b_3h)^2}.$$

Hence the main term we need focus on is

$$\frac{1}{8} \sum_{\substack{\mathbf{d}\in\mathbb{N}^3 \\ \gcd(d_1d_2d_3,2h)=1 \\ \forall\{i,j,k\}=\{1,2,3\}:\gcd(d_i,b_j,b_k)=1 \\ \forall\{i,j,k\}=\{1,2,3\}:d_i\leqslant B/(b_i^2h_{ij}h_{ik})}} \frac{\mu^2(d_1d_2d_3)}{\tau(d_1d_2d_3)}.$$

The lemma now follows from Lemma 6.5.2 □

To complete the analysis of $\mathcal{M}^{(2)}_{\mathbf{b},\mathbf{h},\widetilde{\mathbf{h}}}(B, z_3)$, we compute the remaining sums.

Lemma 6.5.6. *Let*

$$\Sigma(\mathbf{h}) = (-1)^{G(\mathbf{h})/4} \sum_{\substack{\boldsymbol{\epsilon}\in\{1,3\}^3 \\ \widetilde{\mathbf{d}}\in\{\pm1\}^3\setminus\{(1,1,1),(-1,-1,-1)\}}} (-1)^{G_{\mathbf{h}}(\boldsymbol{\epsilon})/4}\left(\frac{-\widetilde{d_2}\widetilde{d_3}}{\epsilon_1}\right)\left(\frac{-\widetilde{d_1}\widetilde{d_3}}{\epsilon_2}\right)\left(\frac{-\widetilde{d_1}\widetilde{d_2}}{\epsilon_3}\right).$$

Then for any $\mathbf{h}$ *we have*

$$\Sigma(\mathbf{h}) = 24.$$

Proof. First we note that it is always the case that $\widetilde{d_i} = \widetilde{d_j} = -\widetilde{d_k}$ for some ordering of $\{i,j,k\} = \{1,2,3\}$. Therefore the product of Jacobi symbols can be rewritten always as $\left(\frac{-1}{\epsilon_k h_{ij}}\right)$. Hence we have

$$\Sigma(\mathbf{h}) = 2(-1)^{G(\mathbf{h})/4} \sum_{\boldsymbol{\epsilon}\in\{\pm 1\}^3} (-1)^{G_{\mathbf{h}}(\boldsymbol{\epsilon})/4} \left(\left(\frac{-1}{\epsilon_1 h_{23}}\right) + \left(\frac{-1}{\epsilon_2 h_{13}}\right) + \left(\frac{-1}{\epsilon_3 h_{12}}\right)\right).$$

Suppose that $\boldsymbol{\epsilon} = (1,1,1)$, then the corresponding summand is

$$\left(\frac{-1}{h_{23}}\right) + \left(\frac{-1}{h_{13}}\right) + \left(\frac{-1}{h_{12}}\right).$$

If $\boldsymbol{\epsilon} = (1,1,-1)$ then the corresponding summand is

$$\left(\frac{-1}{h_{13}}\right) + \left(\frac{-1}{h_{23}}\right) - \left(\frac{-1}{h_{12}h_{13}h_{23}}\right).$$

The terms for when $\boldsymbol{\epsilon} = (1,-1,1)$ and $(-1,1,1)$ are similar. When $\boldsymbol{\epsilon} = (1,-1,-1)$ then the summand is

$$\left(\frac{-1}{h_{12}}\right) + \left(\frac{-1}{h_{13}}\right) + \left(\frac{-1}{h_{12}h_{13}h_{23}}\right).$$

The terms for when $\boldsymbol{\epsilon} = (-1,-1,1)$ and $(-1,1,-1)$ are again analogous. Lastly when $\boldsymbol{\epsilon} = (-1,-1,-1)$ then the summand is

$$\left(\frac{-1}{h_{23}}\right) + \left(\frac{-1}{h_{13}}\right) + \left(\frac{-1}{h_{12}}\right).$$

Summing these terms we see that

$$\Sigma(\mathbf{h}) = 12(-1)^{G(\mathbf{h})/4} \left(\left(\frac{-1}{h_{12}}\right) + \left(\frac{-1}{h_{13}}\right) + \left(\frac{-1}{h_{23}}\right) - \left(\frac{-1}{h_{12}h_{13}h_{23}}\right)\right).$$

Now one can check that for any value of $\mathbf{h}$ the expression above is always 24. □

Remark. Combining Lemmas 6.5.1, 6.5.2, 6.5.4, 6.5.5 and 6.5.6 we can produce an asymptotic for the number of soluble diagonal ternary quadratic forms whose coefficients live in a box of sidelength B and whose coefficients are squarefree and pairwise coprime. This corresponds to setting b_i and m_{ij} equal to 1 in all our computations which yields a main term of

$$\frac{B^3}{(\log B)^{3/2}\Gamma\left(\frac{1}{2}\right)^3} \left(\frac{4}{7}\times 6 \times \frac{5}{2} + \frac{1}{14}\times 24\right) \prod_p \left(1-\frac{1}{p}\right)^{\frac{3}{2}} \left(1+\frac{3}{2p}\right).$$

This is precisely what was obtained by Guo [53, Theorem 1.1].

Finally we can complete our analysis.

Lemma 6.5.7. *For $z_i \leqslant (\log B)^{30}$, we have*

$$\sum_{\substack{\mathbf{b}\in(\mathbb{N}\cap[1,z_1])^3\\ \gcd(b_1,b_2,b_3)=1}} \sum_{\substack{\mathbf{h}=(h_{12},h_{13},h_{23})\in\mathbb{N}^3\\ \widetilde{\mathbf{h}}=(\widetilde{h_{12}},\widetilde{h_{13}},\widetilde{h_{23}})\in\mathbb{N}^3\\ (6.2.5)}} \frac{(-1)^{\frac{G(\mathbf{h})}{4}}}{\tau_{odd}(h\widetilde{h})}\mathcal{M}^{(2)}_{\mathbf{b},\mathbf{h},\widetilde{\mathbf{h}}}(B,z_3)$$

$$= \frac{3}{2\Gamma\left(\frac{1}{2}\right)^3}\frac{B^3}{(\log B)^{\frac{3}{2}}}\prod_p\left(1-\frac{1}{p}\right)^{\frac{3}{2}}\frac{(p^2+p+1)(2p^2+p+2)}{2(p^2-1)^2}$$

$$+ O\left(\frac{B^3(\log\log B)^{\frac{3}{2}}}{(\log B)^{\frac{5}{2}}}\right).$$

Proof. Combining Lemmas 6.5.4 and 6.5.5, it's clear that we just need to compute the sums over $\mathbf{b}$ and $\mathbf{h}$. Again we must relate this sum to the quantity $\mathcal{J}_0$. Since the h_{ij} are all odd we conclude that the sum over the $\mathbf{b}$ and $\mathbf{h}$ is equal to

$$\mathcal{J}_0 \sum_{\substack{\beta_1,\beta_2,\beta_3\in\mathbb{Z}_{\geqslant 0}\\ \min\{\beta_1,\beta_2,\beta_3\}=0}} \frac{1}{4^{\beta_1+\beta_2+\beta_3}} = \frac{7}{3}\mathcal{J}_0.$$

Multiplying this by the $\frac{3}{8}$ incurred by extending the Euler product $\mathcal{J}_0$ from odd primes to all primes, the 24 in Lemma 6.5.6 and the main term in Lemma 6.5.5 the result follows. □

From this and Lemma 6.5.3, we have shown the following.

Theorem 6.5.8. *For any $B \geqslant 1$, we have*

$$N_{loc}(X,B,\pi) = \frac{21}{4\Gamma\left(\frac{1}{2}\right)^3}\frac{B^3}{(\log B)^{3/2}}\prod_p\left(1-\frac{1}{p}\right)^{3/2}\frac{(p^2+p+1)(2p^2+p+2)}{2(p^2-1)^2}$$

$$+ O\left(\frac{B^3(\log\log B)^{5/2}}{(\log B)^{3/2}}\right).$$

6.6 Interpretation of the leading constant

In this final section, we will express the Euler product appearing in Theorem 6.5.8 as a product of local densities. Following the approach of [9] and [15], for each place ν of $\mathbb{Q}$ we will compute σ_ν, the proportion of diagonal ternary conics over $\mathbb{Q}_\nu$ which are isotropic.

6.6.1 $\mathbb{Q}_p$ densities

Let p be an odd prime and denote the usual Haar measure on $\mathbb{Z}_p$ by μ_p. By abuse of notation we denote by μ_p the Haar measure on $\mathbb{Z}_p^2$ and $\mathbb{Z}_p^3$. Let

$$S_{\mathbf{a}} : a_0x_0^2 + a_1x_1^2 + a_2x_2^2 = 0.$$

We shall calculate explicitly

$$\widetilde{\sigma_p} = \mu_p\left(\mathbf{a} \in \mathbb{Z}_p^3 : S_{\mathbf{a}}(\mathbb{Q}_p) \neq \varnothing\right).$$

The proportion of points in $\mathbb{P}^2(\mathbb{Q}_p)$ above which the fibre has a $\mathbb{Q}_p$ point is defined to be

$$\sigma_p := \left(1 - \frac{1}{p}\right)^3 \widetilde{\sigma_p}.$$

First note that the contribution of those $\mathbf{a}$ such that there exists i with $a_i = 0$ is 0 because $\mu_p(\mathbf{a} \in \mathbb{Z}_p^3 : a_1 = 0) = 0$. Thus $\widetilde{\sigma_p} = \sigma_p^*$, where

$$\sigma_p^* := \mu_p\left(\mathbf{a} \in (\mathbb{Z}_p \setminus \{0\})^3 : S_{\mathbf{a}}(\mathbb{Q}_p) \neq \varnothing\right) \tag{6.6.1}$$

Now writing $c_i := p^{-v_p(a_i)} a_i$ and absorbing squares we obtain

$$\sigma_p^* = \sum_{\mathbf{t} \in \{0,1\}^3} \sum_{\substack{\mathbf{v} \in (\mathbb{Z} \cap [0,\infty))^3 \\ \forall i : v_i \equiv t_i \bmod 2}} \mu_p\left(p^{\boldsymbol{\nu}}\mathbf{c} \in \mathbb{Z}_p^3 : \mathbf{c} \in (\mathbb{Z}_p^*)^3 \text{ and } S_{p^{\mathbf{t}}\mathbf{c}}(\mathbb{Q}_p) \neq \varnothing\right), \tag{6.6.2}$$

where $p^{\boldsymbol{\nu}}\mathbf{c}$ is understood to mean $(p^{v_0}c_0, p^{v_1}c_1, p^{v_2}c_2)$. Note that the conic $S_{p^{\boldsymbol{\nu}}\mathbf{c}}$ is soluble if and only if $S_{p^{(\nu_1+1,\nu_2+1,\nu_3+1)}\mathbf{c}}$ is also soluble, since one can multiply the defining equation by p and absorb squares as necessary. If $\mathbf{t} = \mathbf{0}$ then the associated conic is smooth in $\mathbb{F}_p$, hence the measure in (6.6.2) is

$$\mu_p\left((p^{v_0}c_0, p^{v_1}c_1, p^{v_2}c_2) \in \mathbb{Z}_p^3 : \mathbf{c} \in (\mathbb{Z}_p^*)^3\right) = p^{-v_0-v_1-v_2}\left(\frac{(p-1)}{p}\right)^3$$

The same must therefore be true when $\mathbf{t} = (1,1,1)$. If $\mathbf{t} = (1,0,0)$ then the set whose measure we wish to compute in (6.6.2) contain those $\mathbf{c}$ with $-c_2c_3$ being a quadratic residue modulo p. Its measure therefore equals

$$\begin{aligned}
&p^{-v_0}\frac{p-1}{p}\mu_p\left((p^{v_1}c_1, p^{v_2}c_2) \in \mathbb{Z}_p^2 : \mathbf{c} \in (\mathbb{Z}_p^*)^2, \left(\frac{-c_1c_2}{p}\right) = 1\right) \\
&\quad = p^{-v_0}\frac{p-1}{p}p^{-v_1-v_2}\mu_p\left(\mathbf{c} \in (\mathbb{Z}_p^*)^2 : \left(\frac{-c_1c_2}{p}\right) = 1\right) \\
&\quad = p^{-v_0}\frac{p-1}{p}p^{-v_1-v_2}\frac{1}{2}\left(\frac{p-1}{p}\right)^2.
\end{aligned}$$

The same argument gives the same number in the cases with $\mathbf{t} = (0,1,0)$ and $\mathbf{t} = (0,0,1)$. Therefore the count is the same when $\mathbf{t} = (1,1,0), (1,0,1)$ or $(0,1,1)$.

Combining these estimates we get from (6.6.2) that for all $p \neq 2$ one has

$$\sigma_p^* = \frac{(p-1)^3}{p^3}\sum_{\mathbf{t} \in \{0,1\}^3} c(\mathbf{t}) \prod_{i=1}^{3}\left(\sum_{\substack{v=0 \\ v \equiv t_i \bmod 2}}^{\infty} p^{-v}\right),$$

where $c(\mathbf{t})$ equals 1 if $\mathbf{t} = \mathbf{0}$ or if $\mathbf{t} = (1,1,1)$, while, it equals $\frac{1}{2}$ in all other cases. The following holds for $t = 0$ and $t = 1$,

$$\sum_{\substack{v=0 \\ v \equiv t \bmod 2}}^{\infty} p^{-v} = \frac{1}{p^t} \frac{1}{(1-p^{-2})}$$

and yields

$$\sigma_p^* = \frac{(p-1)^3}{p^3(1-p^{-2})^3} \sum_{\mathbf{t} \in \{0,1\}^3} \frac{c(\mathbf{t})}{p^{t_1+t_2+t_3}}.$$

It is now straightforward to check that the sum over $\mathbf{t}$ equals

$$1 \cdot \left(1 + \frac{1}{p^3}\right) + \frac{1}{2}\left(\frac{3}{p} + \frac{3}{p^2}\right) = 1 + \frac{3}{2p} + \frac{3}{2p^2} + \frac{1}{p^3}.$$

Hence when $p \neq 2$ we see from $\widetilde{\sigma_p} = \sigma_p^*$ that

$$\begin{aligned}
\sigma_p &= \left(1 - \frac{1}{p^3}\right) \frac{(p-1)^3}{p^3(1-p^{-2})^3} \left(1 + \frac{3}{2p} + \frac{3}{2p^2} + \frac{1}{p^3}\right) \\
&= \frac{p^3-1}{p^3} \frac{1}{(1+p^{-1})^3} \left(1 + \frac{3}{2p} + \frac{3}{2p^2} + \frac{1}{p^3}\right) \\
&= \frac{(p^3-1)(2p^2+p+2)}{2(p+1)^2p^3}.
\end{aligned}$$

6.6.2 $\mathbb{R}$ density

For the real density, we use the Lebesgue measure and compute

$$\sigma_\infty = \frac{1}{2}\mathrm{vol}\left(\mathbf{a} \in ([-1,1] \cap \mathbb{R})^3 : S_{\mathbf{a}}(\mathbb{R}) \neq \varnothing\right).$$

(The factor $\frac{1}{2}$ here is to deal with multiplication by -1, since we are interested in points on $\mathbb{P}^2$.) This number equals

$$\begin{aligned}
4 &- \frac{1}{2}\mathrm{vol}\left(\mathbf{a} \in ([-1,1] \cap \mathbb{R})^3 : a_1 > 0, a_2 > 0, a_3 > 0\right) \\
&- \frac{1}{2}\mathrm{vol}\left(\mathbf{a} \in ([-1,1] \cap \mathbb{R})^3 : a_1 < 0, a_2 < 0, a_3 < 0\right),
\end{aligned}$$

which is $4 - 2 \times \frac{1}{2} = 3$.

6.6.3 $\mathbb{Q}_2$ density

The case $p = 2$ is similar to the case of odd primes. We again observe that

$$\sigma_2 = \left(1 - \frac{1}{2^3}\right) \mu_2(\{\mathbf{a} \in (\mathbb{Z}_2 \setminus \{0\})^3 : S_{\mathbf{a}}(\mathbb{Q}_2) \neq \varnothing\}),$$

and that this inner measure is given by (6.6.2). When $\mathbf{t} = (0,0,0)$, the measure appearing in (6.6.2) is given by

$$2^{-\nu_0-\nu_1-\nu_2}\mu_2(\{\mathbf{c} \in (\mathbb{Z}_2^\times)^3 : \left(\frac{-c_1c_3, -c_2c_3}{\mathbb{Q}_2}\right) = +1\}).$$

By Lemma 2.2.2, the Hilbert symbol in this case can be written as

$$(-1)^{\frac{-c_1c_3-1}{2}\frac{-c_2c_3-1}{2}}.$$

We may break into residue classes so that the measure we wish to calculate becomes

$$2^{-\nu_0-\nu_1-\nu_2}\sum_{\substack{\mathbf{d} \bmod 4 \\ \mathbf{d}\in E}}^{*} \mu_2(\{\mathbf{c} \in (\mathbb{Z}_2^\times)^3 : \mathbf{c} \equiv \mathbf{d} \bmod 4\}) = 2^{-\nu_0-\nu_1-\nu_2-6}\#E,$$

where E is the set of coprime residue classes mod 4 such that the associated conic is soluble. One now simply checks that $\#E = 4$.

In the case that $\mathbf{t} = (1,0,0)$ the solubility of the conic $pc_0X_0^2+c_1X_1^2+c_2X_2^2 = 0$ is equivalent to the condition

$$(-1)^{\frac{-c_1c_3-1}{2}\frac{-c_2c_3-1}{2}+\frac{c_2^2c_3^2-1}{8}} = +1. \tag{6.6.3}$$

Hence the measure which appears in (6.6.2) is given by

$$2^{-\nu_0-\nu_1-\nu_2-9}\#E',$$

where E' is the set of residue classes mod 8 such that (6.6.3) holds. One can compute easily that $\#E' = 32$.

Analogoulsy to the case when p was odd, these two cases are enough to deduce the total which is given by

$$\sigma_p = \frac{7}{8}\sum_{\mathbf{t}\in\{0,1\}^3}\sum_{\substack{\mathbf{v}\in(\mathbb{Z}\cap[0,\infty))^3 \\ \forall i: v_i\equiv t_i \bmod 2}} 2^{-\nu_0-\nu_1-\nu_2-4} = \frac{7}{2^7}\sum_{\mathbf{v}\in(\mathbb{Z}\cap[0,\infty))^3} 2^{-\nu_0-\nu_1-\nu_2} = \frac{7}{16}.$$

6.6.4 Comparison of local densities with the leading constant

By the estimate

$$\sigma_p = 1 - \frac{3}{2p} + O\left(\frac{1}{p^2}\right)$$

for odd p, we infer that

$$\frac{\sigma_p}{(1-\frac{1}{p})^{3/2}} = 1 + O\left(\frac{1}{p^2}\right)$$

Hence $\prod_p \frac{\sigma_p}{(1-\frac{1}{p})^{3/2}}$ converges. We now define

$$\tau_p := \frac{\sigma_p}{(1-\frac{1}{p})^{3/2}},$$

and

$$\tau_\infty := \sigma_\infty.$$

We may rewrite

$$\begin{aligned}\tau_p &= \frac{p^3-1}{p^3}\frac{2p^2+p+2}{2(p+1)^2(1-\frac{1}{p})^{3/2}} \\ &= \frac{(p-1)(p^2+p+1)}{p^3}\frac{2p^2+p+2}{2(p+1)^2(1-\frac{1}{p})^{3/2}} \\ &= \left(1-\frac{1}{p}\right)^{3/2}\frac{(p-1)(p^2+p+1)(2p^2+p+2)}{2(p+1)^2(p-1)^3}.\end{aligned}$$

Now we observe that

$$\begin{aligned}\prod_\nu \tau_\nu &= \sigma_\infty \prod_p \left(1-\frac{1}{p}\right)^{-3/2}\sigma_p \\ &= 3\times\frac{7\sqrt{2}}{8}\prod_{p>2}\tau_p \\ &= \frac{9}{4}\prod_p\left(1-\frac{1}{p}\right)^{3/2}\frac{(p^2+p+1)(2p^2+p+2)}{2(p^2-1)^2}.\end{aligned}$$

Hence the Euler product appearing in Theorem 6.5.8 is equal to $\frac{4}{9}\prod_\nu \tau_\nu = \frac{4}{3}\prod_p \tau_p$, which is how we deduce Theorem 6.1.1.

Chapter 7

Weak approximation for a family of quadric surface bundles

7.1 Introduction

7.1.1 Rationality problems

Let X be a projective variety defined over a field k. The simplest such is of course the projective space $\mathbb{P}^N$, but there are various types of varieties which are close to being a projective space:

- X is *k-rational* if there is an open dense subset $V \subset X$ which is k-isomorphic to an open dense subset of some projective space $U \subset \mathbb{P}^N_k$,
- X is *k-stably rational* if $X \times \mathbb{P}^m_k$ is k-rational for some $m \in \mathbb{N}$,
- X is *k-unirational* if there is a dominant k-rational map from some projective space $\mathbb{P}^N_k$ to X,

Understanding the relationship between these notions of rationality is one of the most classical problems in algebraic geometry. Evidently a k-rational variety is k-stably rational. If X is stably rational then there is a k-rational map

$$\mathbb{P}^N_k \dashrightarrow X \times \mathbb{P}^m_k \to X,$$

by projecting onto X, hence X is k-unirational. Hence we have the chain of implications

$$k\text{-rational} \implies k\text{-stably rational} \implies k\text{-unirational}$$

In 1876, Lüroth [80] showed that for a smooth projective curve all of these notions are in fact equivalent and equivalent to $X \cong_k \mathbb{P}^1_k$. For surfaces, Castelnuovo and Enriques showed that all the above notions are equivalent when k is

an algebraically closed field. However if one looks at non-algebraically closed k, then there exist k-unirational non k-rational surfaces (e.g. a minimal cubic surface over a number field k is not k-rational [84, Theorem 33.1] but all smooth cubic hypersurfaces are k-unirational [69, Theorem 1.1]). Since the late 19^{th} century, the Lüroth problem of finding smooth complex nonrational unirational varieties of dimension at least 3 was a major open problem in algebraic geometry. This problem was solved in the 1970's, when 3 independent examples were provided:

- Clemens–Griffiths [25] showed that every smooth cubic threefold is non k-rational.
- Iskovskikh–Manin [67] proved that smooth quartic threefolds are non k-rational. It is known that some, such as the one defined by

  $$x_0^4 + x_1^4 + x_2^4 + x_3^4 + x_0x_4^3 + x_3^3x_4 - 6x_1^2x_2^2 = 0,$$

  are k-unirational (hence provide solutions to the Lüroth problem) however it is still unknown whether or not all are.
- Artin–Mumford [1] constructed a k-unirational non k-stably rational variety (it can be realised as a double cover of $\mathbb{P}^3_{\mathbb{C}}$ ramified along a specific quartic).

Therefore the implication

$$k\text{-stably rational} \implies k\text{-unirational}$$

is sharp. The implication

$$k\text{-rational} \implies k\text{-stably rational}$$

was proven to be sharp by Beauville–Colliot-Thélène–Sansuc–Swinerton-Dyer [7] who showed that the conic bundle defined by

$$0 \neq x^2 - ay^2 = f(t)z^2$$

for f an irreducible cubic with $\text{disc}(f)=a$ is stably rational but not rational.

7.1.2 The work of Hassett–Pirutka–Tschinkel

One could also study such rationality properties in families of varieties. Given a fibration $X \xrightarrow{\pi} B$ of smooth projective complex varieties, if some of the fibres satisfy one of the rationality properties need all of the others? The first counterexample to this was proven recently by Hassett–Pirutka–Tschinkel [59].

Theorem 7.1.1. *There exist smooth families of complex projective fourfolds $X \xrightarrow{\pi} B$ with B connected such that for every $b \in B$ the fibre $\pi^{-1}(b)$ is a quadric surface bundle over $\mathbb{P}^2$ and such that*

1. *A very general fibre is not stably $\mathbb{C}-$rational,*
2. *The set of $b \in B$ such that $\pi^{-1}(b)$ is $\mathbb{C}-$rational is dense in B under the Euclidean topology.*

Here the term "very general" means that the point (defining the fibre) lies outside of a union of countably many proper closed subvarieties of the base. Hence Theorem 7.1.1 says that many of the fibres in this family are rational but many are also not even stably rational, and so we say that rationality is not a deformation invariant property for these fourfolds.

Remark. If each of the fibres in a family $X \xrightarrow{\pi} B$ is, for example, a Brauer–Severi variety or a quadric, then the fibre would be rational if and only if it has local points at all places. Hence in this case the set of rational fibres in 2 above is exactly the set counted by $N_{\mathrm{loc}}(X, B, \pi)$ in the Loughran–Smeets problem. However note that in all such fibrations which we have so far encountered there are singular fibres so these do not constitute counter-examples to the deformation invariance of rationality.

In particular, the fibres in the Hassett–Pirutka–Tschinkel family are biquadratic varieties inside $\mathbb{P}^2 \times \mathbb{P}^3$ which acquire the structure of quadric surface bundles via the projection onto $\mathbb{P}^2$. This result was proven using the specialisation method pioneered by Voisin [120] and later refined by Colliot-Thélène–Pirutka [30] and Schreieder [103]. The upshot of this method is that if one can find a special fibre X_b in the family such that

1. $H^2_{\mathrm{nr}}(\mathbb{C}(X_b), \mathbb{Z}/2\mathbb{Z})$ is non-trivial,
2. X_b admits a desingularisation which is universally CH_0-trivial,

then a very general fibre is not stably rational. In [59] these conditions were shown to be satisfied by the fourfold defined for the equation

$$xyt_1^2 + xzt_2^2 + yzt_3^2 + F(x, y, z)t_4^2 = 0, \tag{7.1.1}$$

where

$$F(x, y, z) = x^2 + y^2 + z^2 - 2(xy + xz + yz).$$

7.1.3 Statement of result

Since the unramified cohomology group (c.f. Section 7.2) used in the specialisation method satisfies

$$H^2_{\mathrm{nr}}(\mathbb{C}(X), \mathbb{Z}/2\mathbb{Z}) = \mathrm{Br}_{\mathbb{C}}(\widetilde{X})[2],$$

where $\widetilde{X}$ is a desingularisation for X (the group on the right hand side is independent of the choice of desingularisation). Hassett–Pirutka–Tschinkel had to explicitly construct a non-trivial element, specifically $(xz, yz)_{\mathbb{C}(X)}$, in the $(\mathbb{C})$ Brauer group of X. This element descends to the Brauer group over $\mathbb{Q}$ and hence it is natural for us to ask if it can obstruct the Hasse principle or weak approximation. In this section, given a number field k we will study weak approximation for fourfolds over k of the form (7.1.1).

Theorem 7.1.2. *Let k be a number field and X/k the projective variety defined by the equation*

$$xyt_1^2 + xzt_2^2 + yzt_3^2 + F(x,y,z)t_4^2 = 0,$$

where F satisfies

$$\begin{cases} F(0,y,z), F(x,0,z), F(x,y,0) \text{ are squares in } k \text{ for any } x,y,z \in k, \\ F(1,0,0) = F(0,1,0) = F(0,0,1) = 1. \end{cases} \tag{7.1.2}$$

If F takes only positive values in all real embeddings then X satisfies weak approximation. Otherwise $X(k_\nu)$ has two connected components for ν real and weak approximation holds in one but may fail in the other.

Remark. The Hassett–Pirutka–Tschinkel example certainly satisfies both conditions (7.1.2).

Remark. We say that if F takes only positive values in all real embeddings then it is *totally positive definite*. It will be important in the proofs of Theorem 7.2.8 and Lemma 7.3.5 to note that the conditions of this theorem ensure that F always takes some positive values in all embeddings so the form can never be totally negative definite.

What we are essentially showing in the above is that weak approximation holds for X *away from* the archimedean places of k. The scenario when F is not a totally positive definite form may at first seem strange but this is precisely the situation for certain Châtelet surfaces and intersections of two quadrics.

Example 7.1.3 ([116])**.** Let U be the affine surface over $\mathbb{Q}$ defined by

$$y^2 + z^2 = (4x-7)(x^2-2) \neq 0,$$

and V any smooth compactification. $U(\mathbb{R})$ has two connected components defined by the inequalities $|x| < \sqrt{2}$ and $x > 7/4$. All the rational points on V are contained within the corresponding second component (see [37, Section 12.5] for a concise discussion of this example).

Example 7.1.4 ([31])**.** Consider the variety defined by the simultaneous vanishing in $\mathbb{P}^5$ of the equations

$$\begin{cases} x_1x_2 = x_3^2 + x_4^2, \\ (x_1 - x_2)(3x_1 - 8x_2) = x_5^2 + x_6^2. \end{cases}$$

There are two real components, defined by the inequalities $x_2/x_1 \geqslant 1$ and $0 \leqslant x_2/x_1 \leqslant 3/8$, and the rational points are restricted to the first component.

The reason that the closure of the Brauer–Manin set (and thus the influence of the Brauer–Manin obstruction on approximation of local points) is restricted to a connected component in these examples is because when the

quotient $\mathrm{Br}(V)/\mathrm{Br}(k)$ is finite the Brauer–Manin set $V(\mathbb{A}_k)^{\mathrm{Br}}$ is a clopen subset of $V(\mathbb{A}_K)$. To see this, let $\alpha \in \mathrm{Br}\, V$ a non-constant class. The obstruction set $V(\mathbb{A}_k)^{\alpha}$ associated to α is defined by $\phi_{\alpha}^{-1}(0)$ where $0 \in \mathbb{Q}/\mathbb{Z}$ and $\phi_{\alpha} : V(\mathbb{A}_k) \to \mathbb{Q}/\mathbb{Z}$ is the map defined in Section 1.2 by $(x_\nu)_\nu \mapsto \sum_\nu \mathrm{inv}_\nu \alpha(x_\nu)$. The quotient $\mathbb{Q}/\mathbb{Z}$ takes the discrete topology, hence $\{0\}$ is both closed and open. Therefore by the continuity of ϕ_α so is $V(\mathbb{A}_k)^{\alpha}$. When $\mathrm{Br}\, V/\mathrm{Br}\, k$ is finite then $V(\mathbb{A}_k)^{\mathrm{Br}}$ is a finite intersection of such sets and hence clopen. We are not able to prove in our setting that $\mathrm{Br}\, V/\mathrm{Br}\, k$ is finite but nevertheless the above discussion should soothe the reader's worries about the strange appearance of the statement of Theorem 7.1.2.

Our proof of Theorem 7.1.2 is a combination of algebro-geometric and analytic techniques. We import some of the cohomological ideas of Hassett–Pirutka–Tschinkel to study the Brauer group of X, which allows us to find a non-trivial element in the Brauer group. This is sufficient to determine when weak approximation is obstructed. To understand when weak approximation is satisfied we use an "analytic fibration method", which is the subject of Section 7.3.

7.2 Non-constant classes in the Brauer group

Throughout this and all subsequent sections we fix a number field k and let X/k be the variety in the statement of Theorem 7.1.2. If we hope to exhibit failures of weak approximation we need to find non-trivial elements in the Brauer group. The total space X is singular, if it were not then the Brauer–Manin obstruction would be empty. Indeed the Brauer group of any smooth biprojective Fano hypersurface of dimension at least 3 contains only constant classes (see e.g. [100, Proposition 2.6]). However our discussion of the Brauer–Manin obstruction in Section 1.2 always made the assumption that the variety was smooth. In [40, Section 8], Colliot-Thélène and Xu observe that to discuss the Brauer–Manin obstruction on singular varieties one must work with a smooth projective birational model. Henceforth, we fix a resolution of singularities $\widetilde{X} \xrightarrow{f} X$. Note that $\widetilde{X}$ naturally has a map to $\mathbb{P}^2$ by composition of the resolution map with the fibration map.

Definition 7.2.1 ([40, Definition 8.2])**.** Let S a finite set of places of k. A singular variety V satisfies *central weak approximation at* S if either of the following equivalent conditions holds:

1. $V_{\mathrm{smooth}}(k)$ is dense in $\prod_{\nu\in S} V_{\mathrm{smooth}}(k_\nu)$,
2. If $\phi : \widetilde{V} \to V$ is a resolution of singularities then $V_{\mathrm{smooth}}(k)$ is dense in $\prod_{\nu\in S} \phi(\widetilde{V}(k_\nu))$.

We note that the second condition is independent of the resolution of singularities chosen. When we say that V satisfies weak approximation we mean that the above holds for any choice of S.

We shall compute $\operatorname{Br}\widetilde{X}$ which is contained within the Brauer group of the generic fibre of the map $\widetilde{X} \to \mathbb{P}^2$, which we denote X_η. This generic fibre is a smooth quadric surface Q over $k(\mathbb{P}^2) =: K$. The Hochschild–Serre spectral sequence tells us that the sequence

$$0 \to \operatorname{Pic} Q \to \operatorname{Pic} \overline{Q}^{\operatorname{Gal}(\overline{K}/K)} \to \operatorname{Br} K \to \operatorname{Br}_1 Q \to \mathrm{H}^1(K, \operatorname{Pic} \overline{Q})$$

is exact (see [37, Prop. 4.3.2] and the remark afterwards for this statement).

Since Q is a quadric surface, $\overline{Q} := Q \otimes_k \overline{k}$ is a ruled rational surface, meaning that $\operatorname{Pic} \overline{Q} = \mathbb{Z}e_1 \oplus \mathbb{Z}e_2$ (by e.g. [55, Chapter V, Proposition 2.3]). By the inflation-restriction sequence (e.g. [37, eq. 1.2]), we have the exact sequence

$$0 \to \mathrm{H}^1(\operatorname{Gal}(L/K), \mathbb{Z}e_1 \oplus \mathbb{Z}e_2) \to \mathrm{H}^1(K, \operatorname{Pic} \overline{Q}) \to \mathrm{H}^1(L, \mathbb{Z}^2)^{\operatorname{Gal}(L/K)} \to 0,$$

where L/K is a quadratic extension over which the quadric Q becomes rational. In fact this final group is trivial since $\mathrm{H}^1(L, \mathbb{Z}^2) = \operatorname{Hom}_{cts}(\operatorname{Gal}(\overline{L}/L), \mathbb{Z}^2) \cong \{1\}$. Therefore if the action of Galois on $\operatorname{Pic} \overline{Q}$ is trivial then $\mathrm{H}^1(K, \operatorname{Pic} \overline{Q})$ is trivial. Otherwise, suppose that $f : \operatorname{Gal}(L/K) \to \mathbb{Z}e_1 \oplus \mathbb{Z}e_2$ is a cocyle and denote by σ the non-trivial element in $\operatorname{Gal}(L/K)$. Suppose that $f(\sigma) = ae_1 + be_2$ then

$$0 = f(\mathrm{id}) = f(\sigma^2) = f(\sigma) + \sigma \cdot f(\sigma) = ae_1 + be_2 + be_1 + ae_2.$$

Therefore $a = -b$ and so

$$f(\sigma) = \sigma \cdot (ae_2) - ae_2,$$

making it a coboundary. Hence $\mathrm{H}^1(\operatorname{Gal}(L/K), \mathbb{Z}e_1 \oplus \mathbb{Z}e_2)$ is trivial and thus so is $\mathrm{H}^1(K, \operatorname{Pic} \overline{Q})$. Therefore the Brauer group of K surjects onto the algebraic Brauer group of X_η. However in fact we have in this case $\operatorname{Br}(Q) = \operatorname{Br}_1(Q)$ since Q becomes rational over a quadratic extension. Hence every element of the Brauer group of $\widetilde{X}$ must be a class in $\operatorname{Br} K$.

Remark. Note that Q is defined over K, the function field. The Brauer group of Q is algebraic over K however over the ground field k it may contain transcendental elements.

We have seen that all classes in $\operatorname{Br}\widetilde{X}$ come from $\operatorname{Br} K$. However when given a class in $\operatorname{Br} K$ how does one decide whether or not it has non-constant image in $\operatorname{Br}\widetilde{X}$? The tool we will use for this is *unramified cohomology*. (Note here we follow the conventions of [59] and not those of [91].) A large part of the novelty of this approach is that one never needs to explicitly compute a smooth model of the singular variety in order to study the Brauer group.

Let V be a smooth variety over k and i a natural number. Associated to each discrete rank one valuation v on $k(V)$ there is a homomorphism (known as the *residue map*) between Galois cohomology groups

$$\partial_v^i : H^i(k(V), \mathbb{Z}/2\mathbb{Z}) \to H^{i-1}(\kappa(v), \mathbb{Z}/2\mathbb{Z}),$$

where $\kappa(v)$ is the residue field of v.

Definition 7.2.2. For V a smooth variety over k and $i \in \mathbb{N}_{\geqslant 1}$, the *$i$-th unramified cohomology group* is defined as

$$H^i_{nr}(k(V)) := \bigcap_v \mathrm{Ker}(\partial^i_v),$$

where the intersection is taken over all discrete rank one valuations of $k(V)$ which are trivial on k.

For smooth projective varieties, one need only look at DVR's associated to codimension 1 points. Hence

$$H^i_{nr}(k(V)) = \bigcap_{x \in V^{(1)}} \mathrm{Ker}(\partial^i_x),$$

with the intersection here running over DVR's associated to codimension one points where $\kappa(x)$ is the residue field of x and ∂^i_x the associated residue map. One important property of these groups is their connection to torsion in the Brauer group.

Proposition 7.2.3 ([91, Prop 3.7])**.** *If V is a smooth projective variety over k then*

$$H^2_{nr}(k(V)) \simeq \mathrm{Br}(V)[2].$$

We will use this to find some 2-torsion elements of the Brauer group which obstruct weak approximation. To perform the residue computations we appeal to the following compatibility result.

Proposition 7.2.4 ([91, Prop 3.4])**.** *Let $A \subset B$ be discrete valuation rings with fields of fractions $M \subset L$ respectively. Let π_A (resp. π_B) be the uniformising parameter of A (resp. B) and $\kappa(A)$ (resp. $\kappa(B)$) the residue field of A (resp. B). Let e be the valuation of π_A in B. Then the following diagram commutes*

$$\begin{array}{ccc} H^i(L, \mathbb{Z}/2\mathbb{Z}) & \xrightarrow{\partial^i_B} & H^{i-1}(\kappa(B), \mathbb{Z}/2\mathbb{Z}) \\ {\scriptstyle Res_{M/L}}\uparrow & & \uparrow{\scriptstyle eRes_{\kappa(A)/\kappa(B)}} \\ H^i(M, \mathbb{Z}/2\mathbb{Z}) & \xrightarrow{\partial^i_A} & H^{i-1}(\kappa(A), \mathbb{Z}/2\mathbb{Z}) \end{array}$$

where $Res_{K/L}$ and $Res_{\kappa(A)/\kappa(B)}$ are the restriction maps in Galois cohomology.

This allows us to compute the residue at a codimension 1 point of $\widetilde{X}$ with codimension 1 image in $\mathbb{P}^2$ by studying the residue along its image. However codimension 1 points in $\widetilde{X}$ could also map down to the generic point of $\mathbb{P}^2$ or a closed point. To handle some of these cases we will have to make use of the following.

Proposition 7.2.5 ([91, Cor 3.12])**.** *Let A be a local ring with field of fractions F and residue field κ. Let Q/κ be a quadric corresponding to the vanishing of the quadratic form $q \simeq \langle 1, -a, -b, abd \rangle$. Let v be a discrete valuation on $F(Q)$ with*

valuation ring B. Assume that $A \subset B$ and that upto multiplication by a square in F, the element d is a unit in A and a square in κ. Let $\alpha = (a, b) \in H^2(F, \mathbb{Z}/2\mathbb{Z})$ and α' its image in $H^2(F(Q), \mathbb{Z}/2\mathbb{Z})$. Then $\partial_v^2(\alpha) = 0$.

Our first step is to show that if x is a point on $\mathbb{P}^2$ such that the fibre X_x is smooth then all residues along this fiber must be trivial.

Lemma 7.2.6. *Suppose $x \in \mathbb{P}^{2(1)}$ or x a closed point of $\mathbb{P}^2$, such that X_x is smooth. Then*

$$\partial_x(\alpha) = 0 \text{ for all } \alpha \in \mathrm{Br}(K).$$

Proof. This proof follows closely the proof of the surjectivity of the map in [91, Theorem 3.17]. Let π be a uniformiser for the local ring $\mathcal{O}_{\mathbb{P}^2,x}$, ν_x the associated valuation and $\kappa(x)$ the residue field. Similarly denote $\kappa(\nu)$ the residue field of the valuation associated to the fiber above x. Since X_x is smooth, we can view it as being defined by the vanishing of a quadratic form over $\mathcal{O}_{\mathbb{P}^2,x}$ whose coefficients each have valuation zero. This means that $\kappa(\nu)$ is the function field of a quadric over $\kappa(x)$ and therefore by [91, Theorem 3.10 (ii)] the map $H^1(\kappa(x), \mathbb{Z}/2\mathbb{Z}) \to H^1(\kappa(\nu), \mathbb{Z}/2\mathbb{Z})$ is injective. Suppose that $\alpha \in \mathrm{Br}(K)$. Then we must have $\partial_v(\alpha) = 0$, hence by the commutativity of the diagram in Proposition 7.2.4, we have $\partial_x(\alpha) = 0$. □

Similarly if the image of a codimension one point in $\widetilde{X}$ is the generic point of $\mathbb{P}^2$ then all residues are trivial and hence again by commutativity we get trivial residues at all elements in $\mathrm{Br}(K)$. The upshot is we can reduce to looking at those points above the ramification locus $\{xyzF(x, y, z) = 0\}$. There are two possible cases to consider:

1. valuations v of $K(Q)$ corresponding to the generic point of C_v some codimension one subset on $\mathbb{P}^2$,
2. valuations v of $K(Q)$ corresponding to closed points of $\mathbb{P}^2$.

We claim that the quaternion algebra $(-xz, -yz)_{k(\mathbb{P}^2)}$ is a non-constant element of $\mathrm{Br}\,\widetilde{X}$. To show this we must compute all residues along all valuations of the form (1) and (2). By Proposition 7.2.4, we have the following commutative diagram

$$\begin{array}{ccc} H^2(K(Q), \mathbb{Z}/2\mathbb{Z}) & \xrightarrow{\partial_v^2} & H^1(\kappa(v), \mathbb{Z}/2\mathbb{Z}) \\ \uparrow & & \uparrow \\ H^2(K, \mathbb{Z}/2\mathbb{Z}) & \xrightarrow{\partial_v^2} & H^1(\kappa(C_v), \mathbb{Z}/2\mathbb{Z}). \end{array}$$

If C_v is not one of the lines $\{x = 0\}, \{y = 0\}$ or $\{z = 0\}$ then the image of $(-xz, -yz)$ along the lower arrow of the diagram is trivial. Hence the residue in $K(Q)$ is trivial. Suppose instead now that $C_v = \{x = 0\}$. In this case the discriminant d is given by $F(0, y, z)$ in the residue field, which is square by (7.1.2). Furthermore since the valuation of F with respect to x is 0, it is indeed

a unit in the associated local ring. Hence by Proposition 7.2.5 we conclude that the residue of $(-xz, -yz)$ along v in $K(Q)$ is trivial. The same reasoning holds for v lying above the generic points of $\{y = 0\}$ and $\{z = 0\}$ similarly.

It remains to understand the residue along v where v lies above a closed point P of $\mathbb{P}^2$. There are 3 possibilities:

1. P does not lie on the locus $\{xyz = 0\}$,
2. P lies on exactly one of the lines $\{x = 0\}$,$\{y = 0\}$ or $\{z = 0\}$,
3. P lies on the intersection of two of these lines.

In the first case, by Lemma 7.2.6 the residue is trivial. In the second case, the image of the discriminant $F(x, y, z)$ in the residue field of a point on the line $\{x = 0\}$, say, is square by (7.1.2). Moreover since x does not divide $F(x, y, z)$, by the second part of (7.1.2), $F(x, y, z)$ must be a unit in the DVR associated to the point. Hence by Proposition 7.2.5, the residue is trivial. The exact same argument handles the third and final case.

To conclude, we have successfully demonstrated that $\operatorname{Br} \widetilde{X}$ contains the transcendental element $(-xz, -yz)_{k(\mathbb{P}^2)}$. This allows us to find sufficient conditions for the existence of a Brauer–Manin obstruction.

Remark. In the above proof we have made pivotal use of the conditions (7.1.2). In geometric language they say that the conic which F defines must be tangent to each of the co-ordinate lines and that it does not vanish at the intersection of any of these lines. One intuitive justification for the above theorem is given by Abhyankar's Lemma. Namely, since the ramification locus of the quaternion algebra $(-xy, -yz)_{k(\mathbb{P}^2)}$ is contained within the ramification locus of $X \to \mathbb{P}^2$, the ramification cancels in $\widetilde{X}$ (see e.g. [106, p. 116] for an example of this phenomenon).

Definition 7.2.7. For an adelic point $(P_\nu)_\nu = (x_\nu, y_\nu, z_\nu; \mathbf{t}_\nu)_\nu \in X(\mathbb{A}_k)$, we introduce the notation $\mathscr{S}((P_\nu)_\nu)$ to denote the set of real places ν of k for which $F(x_\nu, y_\nu, z_\nu) < 0$ and x_ν, y_ν and z_ν all have the same sign under the embedding $k \hookrightarrow \mathbb{R}$ associated to ν.

Theorem 7.2.8. *Let $\mathcal{A} = (-xy, -yz)_{k(X)}$. Then*

$$X(\mathbb{A}_k)^{\mathcal{A}} = \{(P_\nu)_\nu = (x_\nu : y_\nu : z_\nu; \mathbf{t}_\nu) \in X(\mathbb{A}_k) : 2 \mid \#\mathscr{S}((P_\nu)_\nu)\}.$$

Remark. When F is totally positive definite the theorem tells us that there is no Brauer–Manin obstruction associated to $\mathcal{A}$. Similarly we see how in the other case, Brauer–Manin obstructions (due to this element) are possible only in one component of a real embedding.

Proof. By definition

$$X(\mathbb{A}_k)^{\mathcal{A}} = \left\{(x_\nu : y_\nu : z_\nu; \mathbf{t}_\nu) \in X(\mathbb{A}_k) : \prod_\nu \left(\frac{-x_\nu z_\nu, -y_\nu z_\nu}{k_\nu}\right) = +1\right\}.$$

At primes $\mathfrak{p}$ such that $\mathfrak{p} \nmid x_\mathfrak{p} y_\mathfrak{p} z_\mathfrak{p} \mathfrak{o}_k$ we have $\left(\frac{-x_\mathfrak{p} z_\mathfrak{p}, -y_\mathfrak{p} z_\mathfrak{p}}{k_\mathfrak{p}}\right) = +1$. Now suppose that $\mathfrak{p} \mid x_\mathfrak{p} \mathfrak{o}_k$. By (7.1.2), $F(x_\mathfrak{p}, y_\mathfrak{p}, z_\mathfrak{p})$ is a square mod $\mathfrak{p}$. Hence because $(x_\mathfrak{p} : y_\mathfrak{p} : z_\mathfrak{p}; t_\mathfrak{p})$ is a point on X we must, by Theorem 2.2.3, have

$$\left(\frac{-x_\mathfrak{p} z_\mathfrak{p}, -y_\mathfrak{p} z_\mathfrak{p}}{k_\mathfrak{p}}\right)\left(\frac{-x_\mathfrak{p} y_\mathfrak{p}, -y_\mathfrak{p} z_\mathfrak{p}}{k_\mathfrak{p}}\right)\left(\frac{-x_\mathfrak{p} y_\mathfrak{p}, -x_\mathfrak{p} z_\mathfrak{p}}{k_\mathfrak{p}}\right) = +1.$$

It follows from the explicit description of the local Hilbert symbols that each of the three symbols in the product is equal and so

$$\left(\frac{-x_\mathfrak{p} z_\mathfrak{p}, -y_\mathfrak{p} z_\mathfrak{p}}{k_\mathfrak{p}}\right) = +1.$$

Of course the reasoning is the same when $\mathfrak{p} \mid y_\mathfrak{p} z_\mathfrak{p} \mathfrak{o}_k$ and so the result follows. Hence the only non-trivial contributions to the product occur at real places. If $F(x_\nu, y_\nu, z_\nu) > 0$ for a real embedding ν then by continuity $F(x, y, z) > 0$ and hence by Theorem 2.2.3, we have $\left(\frac{-xy, -yz}{k_\nu}\right)\left(\frac{-xy, -xz}{k_\nu}\right)\left(\frac{-xz, -yz}{k_\nu}\right) = +1$. This is only possible when x, y and z do not all have the same sign (in this real embedding) but this immediately implies that $\left(\frac{-xy, -yz}{k_\nu}\right) = +1$. However when $F(x, y, z) < 0$ in a real embedding and F indefinite, it could be possible for x, y, z to all have the same sign and hence for the Hilbert symbol to be -1. □

So far we have demonstrated that there exists a non-trivial class in the Brauer group of $\widetilde{X}$ and given sufficient conditions for this element to obstruct weak approximation. In the final section, we show that if these conditions are not met then weak approximation holds.

7.3 Verifying weak approximation analytically

The definition of weak approximation in Definition 1.1.10 can be re-interpreted as follows. Weak approximation holds for X if for any finite set of places S and any nonempty open sets $U_\nu \subset X(k_\nu)$ for $\nu \in S$, there exists $x \in X(k)$ such that $x \in U_\nu$ for all $\nu \in S$. We can make a simplification reducing the problem to approximating local base points by rational base points with rational fibres.

Lemma 7.3.1. *Weak approximation holds for X if and only if for any finite set of places S, local base points $b_\nu = (x_\nu : y_\nu : z_\nu) \in \mathbb{P}^2(k_\nu)$ and non-empty neighbourhoods $b_\nu \in U_\nu$ for each $\nu \in S$ there exists a rational base point $b = (x : y : z) \in \mathbb{P}^2(k)$ such that $b \in U_\nu$ for each $\nu \in S$ and $\pi^{-1}(b)(k) \neq \varnothing$.*

Proof. If weak approximation holds then certainly the latter condition is true, by just projecting onto $\mathbb{P}^2$. The proof in the opposite direction follows closely the proof of Theorem 2.1.2. □

Let X_b denote the fibre above the rational point $b \in \mathbb{P}^2(k)$. By the Hasse–Minkowski theorem, X_b has rational points if and only if it has local points everywhere. The next result shows that the approximation condition ensures the solubility of fibres.

Lemma 7.3.2. *If $\nu \in S$ then $X_b(k_\nu) \neq \varnothing$.*

Proof. If ν is a complex place then this is trivially true. Otherwise the solubility of the fibre X_b is determined by Theorem 2.2.3. If b_ν is a point on $\mathbb{P}^2$ such that $X_{b_\nu}(k_\nu) \neq \varnothing$ then there exists an open neighbourhood U_ν such that $b \in U_\nu$ implies that $X_b(k_\nu) \neq \varnothing$. Indeed, in the case that ν is a real place, the conditions for solubility are just conditions on the signs of the coefficients of the quadratic form defining the fibre. Since each of these coefficients is a continuous function of the co-ordinates x, y and z, the solubility conditions will be preserved for all points sufficiently close to b_ν. In the non-archimedean case, solubility is determined by the squareness of the discriminant of the quadratic form or by a Hilbert symbol condition, each of which depend only on the residue class of the point modulo a large power of the prime corresponding to ν. Hence, possibly after shrinking ϵ, we deduce that weak approximation ensures solubility at $\nu \in S$. □

Enlarging S as necessary we assume that all archimedean places and all places dividing $2\mathfrak{o}_k$ are in S. Hence we need only investigate the solubility of X_b in $k_\mathfrak{p}$ for places $\mathfrak{p}$ above odd rational primes. The fibre X_b is defined by a diagonal quaternary quadratic form, the solubility of which is determined by Theorem 2.2.3. One sees immediately that the condition $d \neq 1$ is equivalent to asking that $F(x, y, z)$ is not square in $k_\mathfrak{p}$. If it is square then we must investigate the product of Hilbert symbols ϵ which is given by

$$\left(\frac{xy, yz}{k_\mathfrak{p}}\right)\left(\frac{xy, xz}{k_\mathfrak{p}}\right)\left(\frac{yz, xz}{k_\mathfrak{p}}\right)\left(\frac{xy, F(x,y,z)}{k_\mathfrak{p}}\right)\left(\frac{xz, F(x,y,z)}{k_\mathfrak{p}}\right)\left(\frac{yz, F(x,y,z)}{k_\mathfrak{p}}\right).$$

Using Lemma 2.2.1, transforms the equation $\epsilon = \left(\frac{-1,-1}{k_\mathfrak{p}}\right)$ into

$$\left(\frac{xy, yz}{k_\mathfrak{p}}\right)\left(\frac{xy, xz}{k_\mathfrak{p}}\right)\left(\frac{xz, yz}{k_\mathfrak{p}}\right) = +1.$$

7.3.1 Counting fibres

We have seen that to prove weak approximation on X it suffices to be able to find a rational point that arbitrarily well approximates a given finite collection of local points on $\mathbb{P}^2$ and such that the fibre above it has a k-point. We will search for such points by passing to the affine cone of $\mathbb{P}^2$, the k-points of which can be represented by triples $(x, y, z) \in \mathfrak{o}_k^3$ such that $\gcd(x\mathfrak{o}_k, y\mathfrak{o}_k, z\mathfrak{o}_k) = \mathfrak{o}_k$. Note that such a representative is unique upto multiplication by units. The $k_\mathfrak{p}$ points on the affine cone are described similarly. We re-interpret the approximation condition at finite primes as follows: Let S_{fin} be a finite set of primes of k, fix $\epsilon > 0$ and $(x_\mathfrak{p}, y_\mathfrak{p}, z_\mathfrak{p})_\mathfrak{p}$ a collection of representatives for some choice of local points in the affine cone of $\mathbb{P}^2$ for each $\mathfrak{p} \in S_{\text{fin}}$. The weak approximation condition asks that one can find a representative (x, y, z) for a rational point in

the affine cone such that

$$|x - x_{\mathfrak{p}}|_{\mathfrak{p}} < \epsilon,$$
$$|y - y_{\mathfrak{p}}|_{\mathfrak{p}} < \epsilon,$$
$$|z - z_{\mathfrak{p}}|_{\mathfrak{p}} < \epsilon,$$

for each $\mathfrak{p} \in S_{\text{fin}}$. Note that these inequalities are unaffected by multiplication by a unit and hence do not depend on the choice of representative for the point in the affine cone. Using weak approximation on $\mathbb{P}^2$ and the Chinese Remainder Theorem, we can find $\mathbf{a} \in \mathfrak{o}_k^3$ such that

$$|x_{\mathfrak{p}} - a_1|_{\mathfrak{p}} < \epsilon, |y_{\mathfrak{p}} - a_2|_{\mathfrak{p}} < \epsilon \text{ and } |z_{\mathfrak{p}} - a_3|_{\mathfrak{p}} < \epsilon,$$

for all finite $\mathfrak{p} \in S$. Therefore the weak approximation condition at the finite places is reinterpreted as the congruence condition

$$(x, y, z) \equiv \mathbf{a} \bmod \mathfrak{m}, \tag{7.3.1}$$

where $\mathfrak{m} = \prod_{\mathfrak{p} \in S} \mathfrak{p}^{m_{\mathfrak{p}}}$, for some suitably large exponents $m_{\mathfrak{p}}$.

It remains to reinterpret the weak approximation condition at the infinite places. Given a fixed $\epsilon > 0$ and a collection of points $(x_\nu : y_\nu : z_\nu) \in \mathbb{P}^2(k_\nu)$ at each infinite place ν, we aim to find a rational point $(x : y : z) \in \mathbb{P}^2(k)$ which lies inside an ϵ neighbourhood of each $(x_\nu : y_\nu : z_\nu)$. We make some choice of representatives $(x_\nu, y_\nu, z_\nu) \in k_\nu^3$ and then the weak approximation condition asks that there exists $(x, y, z) \in \mathfrak{o}_k^3$ such that $\gcd(x\mathfrak{o}_k, y\mathfrak{o}_k, z\mathfrak{o}_k) = \mathfrak{o}_k$ and $\exists \lambda_\nu \in k_\nu$ for each ν such that

$$|\lambda_\nu x_\nu - x|_\nu < \epsilon$$
$$|\lambda_\nu y_\nu - y|_\nu < \epsilon$$
$$|\lambda_\nu z_\nu - z|_\nu < \epsilon.$$

This condition is independent of the choice of representatives for the local points because we may adjust the scaling parameters λ_ν accordingly. Ultimately we would like to be able to consider the conditions at each of the infinite places simultaneously. To do this we follow the approach of Skinner [110, Section 5]. Let $\tau_1, \ldots, \tau_d$ a $\mathbb{Z}$-basis for $\mathfrak{o}_k$. k has r real embeddings, denoted $\sigma_1, \ldots, \sigma_r$, and $2s$ complex embeddings, denoted $\sigma_{r+1}, \ldots, \sigma_{r+2s}$, where $\sigma_{r+s+i} = \overline{\sigma_{r+i}}$ for $i = 1, \ldots, s$. Denote by k_i the completion of k with respect to the embedding σ_i. We may identify the direct sum of these completions with the commutative $\mathbb{R}$-algebra

$$V := k \otimes_{\mathbb{Q}} \mathbb{R} = \bigoplus_{i=1}^{r+s} k_i.$$

There is a canonical embedding of k into V given by $x \mapsto \oplus \sigma_i(x)$. We identify k with its image under this embedding. Doing so realises $\mathfrak{o}_k$ as a lattice in V and $\tau_1, \ldots, \tau_d$ as a real basis for $\mathfrak{o}_k$. We define a distance on V with respect to this basis by

$$|x|_\tau = |x_1\tau_1 + \cdots + x_d\tau_d|_\tau := \max_i |x_i|. \tag{7.3.2}$$

Note that since projection on to the i^{th} co-ordinate is linear, there exists a constant c, depending at most on k and τ, such that

$$|\sigma_i(x)| \leqslant c|x|_\tau. \tag{7.3.3}$$

This can be extended to a distance on V^n. If $\mathbf{x} = (x^{(1)}, \dots, x^{(n)}) \in V^n$ then $|\mathbf{x}|_\tau = \max_j |x^{(j)}|_\tau$. The usual field norm on k extends to V via $N(x) = \sigma_1(x) \dots \sigma_r(x)|\sigma_{r+1}(x)|^2 \dots |\sigma_{r+s}(x)|^2$. Throughout (as above) a superscript index will always refer to the component of a vector and a subscript index will refer to the co-ordinate with respect to a fixed basis.

Given a point $\mathbf{b} \in V^n$ we denote the ball of radius ρ centred on $\mathbf{b}$ by $\mathcal{B}(\mathbf{b}, \rho) := \{\mathbf{x} \in V^n : |\mathbf{x}-\mathbf{b}|_\tau < \rho\}$. Now we fix $\epsilon > 0$ and make some choice of S a finite set of places (including all archimedean places) and local points (x_ν, y_ν, z_ν) in the affine cone for each $\nu \in S$. Let $D \in \mathbb{Z}$ satisfying $D \equiv 1 \bmod \mathfrak{m}$ and such that $D > \frac{c}{\epsilon}$. Then for each infinite place ν_i, with corresponding embedding σ_i, we define

$$(q_{\nu_i}, r_{\nu_i}, s_{\nu_i}) = D(x_{\nu_i}, y_{\nu_i}, z_{\nu_i}) \in k_i^3.$$

Combining these local points gives a point in V^3 by the above identification

$$\boldsymbol{\zeta}_0 = (\zeta_0^{(1)}, \zeta_0^{(2)}, \zeta_0^{(3)}) := \bigoplus_{i=1}^{r+s} (q_{\nu_i}, r_{\nu_i}, s_{\nu_i}).$$

Let $\mathcal{R} = \mathcal{B}(\boldsymbol{\zeta}_0, \rho)$ for some $\rho < 1$. If (x, y, z) satisfies (7.3.1) and $(x, y, z) \in \mathcal{R}$ then $(D^{-1}x, D^{-1}y, D^{-1}z)$ lies inside an ϵ neighbourhood of (x_ν, y_ν, z_ν) at all places $\nu \in S$. Indeed at the infinite place ν_i, we have

$$\begin{aligned} |D^{-1}x - x_{\nu_i}|_{\nu_i} &\leqslant D^{-1}|x - q_{\nu+i}|_{\nu_i} \\ &< cD^{-1}|x - \zeta_0^{(1)}|_\tau \\ &< \epsilon. \end{aligned}$$

The parameter D will be fixed and ensures that every point in $\mathcal{R}$ satisfying the congruence condition also satisfies the weak approximation conditions. However it might not be the case that such a point necessarily exists. We introduce one more parameter $P \equiv 1 \bmod \mathfrak{m}$ which we will send to infinity and we now search for points in the expanding region $(x, y, z) \in P\mathcal{R}$ satisfying (7.3.1). This will yield points $(D^{-1}P^{-1}x, D^{-1}P^{-1}y, D^{-1}P^{-1}z)$ which arbitrarily well approximate our chosen representatives for the given local points in the affine cone.

We now introduce the following set, a non-trivial lower bound for which will suffice to prove weak approximation for X,

$$\mathcal{A} := \{(x, y, z) \in \mathfrak{o}_k^3 : (x, y, z) \equiv \mathbf{a} \bmod \mathfrak{m}, (x, y, z) \in P\mathcal{R} \text{ and } X_{(x:y:z)}(\mathbb{Q}) \neq \varnothing\}.$$

Let $\mathfrak{d}_i = \gcd(a_i\mathfrak{o}_k, \mathfrak{m})$. To count points b in $\mathcal{A}$ we need to ensure that X_b is soluble at all primes dividing $xyz\mathfrak{o}_k$ that are not in S. To simplify this count we search only for points such that

$$x\mathfrak{o}_k = \mathfrak{d}_1\mathfrak{q}_1, y\mathfrak{o}_k = \mathfrak{d}_2\mathfrak{q}_2 \text{ and } z\mathfrak{o}_k = \mathfrak{d}_3\mathfrak{q}_3$$

for $\mathfrak{q}_1, \mathfrak{q}_2, \mathfrak{q}_3$ distinct prime ideals. Further we assume that the norm of each of the ideals $\mathfrak{q}_i$ is larger than that of any prime in S. Then we need only check the solubility of X_b at the primes $\mathfrak{q}_i$. We saw previously that to establish solubility at a prime $\mathfrak{p}$ it is sufficient to prove

$$\left(\frac{xy, yz}{k_{\mathfrak{p}}}\right)\left(\frac{xy, xz}{k_{\mathfrak{p}}}\right)\left(\frac{xz, yz}{k_{\mathfrak{p}}}\right) = +1,$$

since then the conditions of Theorem 2.2.3 are satisfied regardless of whether F is a square in $K_{\mathfrak{p}}$ or not. Recalling Lemma 2.2.2, we see that the condition for solubility at $\mathfrak{q}_1$ is

$$\left(\frac{-yz}{\mathfrak{q}_1}\right) = +1.$$

Similarly, at $\mathfrak{q}_2$ and $\mathfrak{q}_3$ we have the conditions

$$\left(\frac{-xz}{\mathfrak{q}_2}\right) = +1 = \left(\frac{-xy}{\mathfrak{q}_3}\right).$$

Therefore,

$$\begin{aligned}\#\mathscr{A} &\geqslant \frac{1}{2} \sum_{\substack{(x,y,z)\in\mathfrak{o}_k^3\cap P\mathcal{R}\\ \gcd(x\mathfrak{o}_k, y\mathfrak{o}_k, z\mathfrak{o}_k)=1\\ xyzF(x,y,z)\neq 0\\ (x,y,z)\equiv \mathbf{a} \bmod \mathfrak{m}\\ x\mathfrak{o}_k=\mathfrak{d}_1\mathfrak{q}_1, y\mathfrak{o}_k=\mathfrak{d}_2\mathfrak{q}_2, z\mathfrak{o}_k=\mathfrak{d}_3\mathfrak{q}_3\\ \mathfrak{q}_i \text{ distinct primes}}} \mathbf{1}_{X_{(x,y,z)}(k)\neq\varnothing} \\ &= \frac{1}{16} \sum_{\substack{(x,y,z)\in\mathfrak{o}_k^3\cap P\mathcal{R}\\ \gcd(x\mathfrak{o}_k, y\mathfrak{o}_k, z\mathfrak{o}_k)=1\\ (x,y,z)\equiv \mathbf{a} \bmod \mathfrak{m}\\ x\mathfrak{o}_k=\mathfrak{d}_1\mathfrak{q}_1, y\mathfrak{o}_k=\mathfrak{d}_2\mathfrak{q}_2, z\mathfrak{o}_k=\mathfrak{d}_3\mathfrak{q}_3\\ \mathfrak{q}_i \text{ distinct primes}}} \left(1+\left(\frac{-xy}{\mathfrak{q}_3}\right)\right)\left(1+\left(\frac{-xz}{\mathfrak{q}_2}\right)\right)\left(1+\left(\frac{-yz}{\mathfrak{q}_1}\right)\right) + O(P^2).\end{aligned}$$

We can expand the brackets into four kinds of sums

$$\begin{aligned}&\sum_{x,y,z} 1 \\ &+ \sum_{x,y,z} \left(\left(\frac{-yz}{\mathfrak{q}_1}\right) + \left(\frac{-xz}{\mathfrak{q}_2}\right) + \left(\frac{-xy}{\mathfrak{q}_3}\right)\right) \\ &+ \sum_{x,y,z} \left(\left(\frac{-xy}{\mathfrak{q}_3}\right)\left(\frac{-yz}{\mathfrak{q}_1}\right) + \left(\frac{-xy}{\mathfrak{q}_3}\right)\left(\frac{-xz}{\mathfrak{q}_2}\right) + \left(\frac{-xz}{\mathfrak{q}_2}\right)\left(\frac{-yz}{\mathfrak{q}_1}\right)\right) \\ &+ \sum_{x,y,z} \left(\frac{-yz}{\mathfrak{q}_1}\right)\left(\frac{-xz}{\mathfrak{q}_2}\right)\left(\frac{-xy}{\mathfrak{q}_3}\right).\end{aligned}$$

7.3.2 The error term

In all but the first and last sums above, there exists one of the $\mathfrak{q}_i$ that divides one of the ideals defined by the variables in the numerator of the Jacobi symbol but does not appear in the denominator. For example, in the term $\left(\frac{-xy}{\mathfrak{q}_3}\right)\left(\frac{-xz}{\mathfrak{q}_2}\right)$, $\mathfrak{q}_1$ divides $x\mathfrak{o}_k$ but does not appear in the denominator. We can exploit cancellation in the sum over terms like this by using a number field extension of Heath-Brown's large sieve for quadratic characters due to Goldmakher–Louvel [52, Theorem 1.1].

Lemma 7.3.3. *Let $\mathfrak{c}$ be an integral ideal of k and $\{\chi_{\mathfrak{a}}\}$ a quadratic Hecke family with respect to $\mathfrak{c}$. Fix any $\epsilon > 0, M, N \geqslant 1$ and a sequence $(\lambda_{\mathfrak{b}})$ of complex numbers parametrised by integral ideals of k. Then, we have*

$$\sum_{N(\mathfrak{a})\leqslant M}^{*}\left|\sum_{N(\mathfrak{b})\leqslant N}^{*}\lambda_{\mathfrak{b}}\chi_{\mathfrak{b}}(\mathfrak{a})\right|^2 \ll_{k,\mathfrak{c},\epsilon} (MN)^{\epsilon}(M+N)\sum_{N(\mathfrak{b})\leqslant N}^{*}|\lambda_{\mathfrak{b}}|^2,$$

where $\sum^{*}$ indicates summation over squarefree ideals $\equiv 1 \bmod \mathfrak{c}$.

We refer the reader to the original paper for the definition of a Hecke family. All that is relevant for our purposes is that the notion generalises the norm residue symbol. By which we mean that there exists a Hecke familiy such that $\chi_{(a)}(\mathfrak{b}) = \left(\frac{a}{\mathfrak{b}}\right)$. We shall use this in the following bilinear form.

Lemma 7.3.4. *Let $\mathfrak{c}$ be an integral ideal in k. If $\alpha_{\mathfrak{a}}$,$\beta_{\mathfrak{b}}$ are sequences bounded in modulus by 1, supported on algebraic integers $\equiv 1 \bmod \mathfrak{c}$ then we have*

$$\sum_{N(\mathfrak{a})\leqslant M}\sum_{N(\mathfrak{b})\leqslant N}^{*}\alpha_{\mathfrak{a}}\beta_{\mathfrak{b}}\chi_{\mathfrak{b}}(\mathfrak{a}) \ll_{\epsilon} (MN)^{1+\epsilon}\left(M^{-\frac{1}{2}}+N^{-\frac{1}{2}}\right).$$

Proof. First we'll use the previous lemma to prove the bilinear statement under the assumption that $\mathfrak{a}$ is squarefree. An application of Cauchy–Schwarz yields

$$\begin{aligned}
\sum_{N(\mathfrak{a})\leqslant M}^{*}\sum_{N(\mathfrak{b})\leqslant N}^{*}\alpha_{\mathfrak{a}}\beta_{\mathfrak{b}}\chi_{\mathfrak{b}}(\mathfrak{a}) &\ll \sum_{N(\mathfrak{a})\leqslant M}^{*}\left|\sum_{N(\mathfrak{b})\leqslant N}^{*}\beta_{\mathfrak{b}}\chi_{\mathfrak{b}}(\mathfrak{a})\right| \\
&\ll M^{\frac{1}{2}}\left(\sum_{N(\mathfrak{a})\leqslant M}^{*}\left|\sum_{N(\mathfrak{b})\leqslant N}^{*}\beta_{\mathfrak{b}}\chi_{\mathfrak{b}}(\mathfrak{a})\right|^2\right)^{\frac{1}{2}} \\
&\ll M^{\frac{1}{2}}\left((MN)^{\epsilon}(M+N)\sum_{N(\mathfrak{b})\leqslant N}^{*}|\beta_{\mathfrak{b}}|^2\right)^{\frac{1}{2}} \\
&\ll M^{\frac{1}{2}}\left(M^{\epsilon}N^{1+\epsilon}(M+N)\right)^{\frac{1}{2}}.
\end{aligned}$$

In general we have

$$\sum_{N(\mathfrak{a})\leqslant MN(\mathfrak{b})\leqslant N}^{*}\alpha_{\mathfrak{a}}\beta_{\mathfrak{b}}\chi_{\mathfrak{b}}(\mathfrak{a}) = \sum_{N(\mathfrak{d})\leqslant M^{\frac{1}{2}}}\sum_{N(\mathfrak{e})\leqslant M/N(\mathfrak{d})^2}^{*}\sum_{N(\mathfrak{b})\leqslant N}^{*}\alpha_{\mathfrak{e}\mathfrak{d}^2}\beta_{\mathfrak{b}}\chi_{\mathfrak{b}}(\mathfrak{a})$$
$$\ll \sum_{N(\mathfrak{d})\leqslant M^{\frac{1}{2}}}(MN)^{\epsilon}\left(\frac{M}{N(\mathfrak{d})^2}N^{\frac{1}{2}}+\left(\frac{M}{N(\mathfrak{d})^2}\right)^{\frac{1}{2}}N\right).$$

□

We'll apply this, for example, to

$$\sum_{x,y,z}\left(\frac{-xy}{\mathfrak{q}_3}\right)\left(\frac{-xz}{\mathfrak{q}_2}\right).$$

First, we observe how the condition $(x,y,z)\in P\mathcal{R}$ restricts the norm of $x\mathfrak{o}_k$. By (7.3.3), we have $N(x)\asymp P^d$ for any $x\in P\mathcal{R}$. Let $\alpha_{x\mathfrak{o}_k}$ be the product of any remaining Legendre symbols containing x and an indicator function for principal ideals that have a generator x in the correct interval which is congruent to $a \bmod \mathfrak{m}$ and such that $x\mathfrak{d}_1^{-1}$ is prime. Similarly define $\beta_{\mathfrak{q}_2}$ to encapsulate the conditions that $\mathfrak{q}_2$ is prime and all the necessary conditions on $\mathfrak{q}_2\mathfrak{d}_2$. Then Lemma 7.3.4 yields

$$\sum_{x,y,z}\left(\frac{-xy}{\mathfrak{q}_3}\right)\left(\frac{-xz}{\mathfrak{q}_2}\right) \ll \sum_{N(z)\ll P^d}\left|\sum_{\substack{N(x)\ll P^d\\ N(\mathfrak{q}_2)\ll P^d/N(\mathfrak{d}_2)}}\alpha_{x\mathfrak{o}_K}\beta_{\mathfrak{q}_2}\left(\frac{x}{\mathfrak{q}_2}\right)\right|$$
$$\ll_{\epsilon} P^{d(3-\frac{1}{2})+\epsilon}.$$

7.3.3 The main term

This leaves us to investigate

$$\sum_{x,y,z}\left[1+\left(\frac{-yz}{\mathfrak{q}_1}\right)\left(\frac{-xz}{\mathfrak{q}_2}\right)\left(\frac{-xy}{\mathfrak{q}_3}\right)\right].$$

The sum $\sum_{x,y,z}1$ has size $\gg P^{3d}/(\log P)^3$ (a fact which will be shown later). Hence proving that $\mathscr{A}$ is non-empty would follow from showing that the above sum does not exhibit too much cancellation. In the following lemma, we will show that if $b_\nu=(x_\nu:y_\nu:z_\nu)$ is the projection of a point belonging to the Brauer–Manin set from the previous section then there is no cancellation in the sum. Recall that for an adelic point $(P_\nu)_\nu=(x_\nu:y_\nu:z_\nu;\mathbf{t}_\nu)_\nu\in X(\mathbb{A}_k)$ we defined the set $\mathscr{S}((P_\nu)_\nu)$ to be those real places ν where $F(x_\nu,y_\nu,z_\nu)<0$ in the corresponding real embedding and x_ν,y_ν and z_ν all have the same sign.

Lemma 7.3.5. *Suppose that* $(x_\nu:y_\nu:z_\nu)_\nu\in\mathbb{P}^2(\mathbb{A}_k)$ *such that*

- *each* $X_{(x_\nu:y_\nu:z_\nu)}(k)$ *is non-empty for every* $\nu \mid \infty$,
- $2 \mid \#\mathscr{S}((P_\nu)_\nu)$ *for any* $(P_\nu)_\nu = (x_\nu, y_\nu, z_\nu; \mathfrak{t}_\nu)_\nu \in X(\mathbb{A}_k)$.

Then for any $(x : y : z) \in \mathbb{P}^2(k)$ *living simultaneously in a small* ν*-adic neighbourhood of* $(x_\nu : y_\nu : z_\nu)$ *for each* $\nu \in S$, *we have*

$$\left(\frac{-yz}{\mathfrak{q}_1}\right)\left(\frac{-xz}{\mathfrak{q}_2}\right)\left(\frac{-xy}{\mathfrak{q}_3}\right) = 1.$$

Otherwise the product above equals -1.

Proof. By the Hilbert product formula

$$\prod_\nu \left(\frac{-xy, -yz}{k_\nu}\right) = +1.$$

Of course $\left(\frac{-xy,-yz}{k_\mathfrak{p}}\right) = +1$ for any $\mathfrak{p} \nmid xyz\mathfrak{o}_k$. By the same reasoning as in the proof of Theorem 7.2.8, if $\mathfrak{p} \in S$ then $\left(\frac{-x_\mathfrak{p} y_\mathfrak{p}, -y_\mathfrak{p} z_\mathfrak{p}}{k_\mathfrak{p}}\right) = +1$. Hence the weak approximation congruence (7.3.1) means that $\left(\frac{-xy,-yz}{k_\mathfrak{p}}\right) = +1$ for each $\mathfrak{p} \in S$. Therefore

$$\prod_{i=1}^{3} \left(\frac{-xy, -yz}{k_{\mathfrak{q}_i}}\right) \prod_{\nu|\infty} \left(\frac{-xy, -yz}{k_\nu}\right) = +1.$$

Applying the explicit description of the local Hilbert symbols gives

$$\prod_{i=1}^{3} \left(\frac{-xy, -yz}{k_{\mathfrak{q}_i}}\right) = \left(\frac{-yz}{\mathfrak{q}_1}\right)\left(\frac{-xz}{\mathfrak{q}_2}\right)\left(\frac{-xy}{\mathfrak{q}_3}\right)$$

and thus

$$\left(\frac{-yz}{\mathfrak{q}_1}\right)\left(\frac{-xz}{\mathfrak{q}_2}\right)\left(\frac{-xy}{\mathfrak{q}_3}\right) = \prod_{\nu|\infty} \left(\frac{-xy, -yz}{k_\nu}\right).$$

This is precisely the product of Hilbert symbols which defines the obstruction set associated to the Brauer class $\mathcal{A}$ demonstrated in the previous section. □

Here we see how the dichotomy of outcomes in Theorem 7.1.2, depending on the definiteness of F, is manifested. If F is totally positive definte then any local points satisfy the conditions of the lemma because the set $\mathscr{S}((P_\nu)_\nu)$ will always be empty. However when F is not totally positive definite, the value of the product of characters depends on the parity of F at the real embeddings. Henceforth we shall assume that the local points satisfy the conditions of Lemma 7.3.5.

All that's left is to show that the counting function $N(P)$ is significantly larger than the error term arising from the application of the large sieve for

quadratic characters, where

$$N(P) := \sum_{\substack{(x,y,z)\in \mathfrak{o}_k^3\cap P\mathcal{R}\\ \gcd(x\mathfrak{o}_k,y\mathfrak{o}_k,z\mathfrak{o}_k)=1\\ (x,y,z)\equiv \mathfrak{a} \bmod \mathfrak{m}\\ x\mathfrak{o}_k=\mathfrak{d}_1\mathfrak{q}_1, y\mathfrak{o}_k=\mathfrak{d}_2\mathfrak{q}_2, z\mathfrak{o}_k=\mathfrak{d}_3\mathfrak{q}_3\\ \mathfrak{q}_i \text{ distinct primes}}} 1.$$

Note that since the $\mathfrak{q}_i$ are distinct and $(x_\mathfrak{p} : y_\mathfrak{p} : z_\mathfrak{p}) \in \mathbb{P}^2(k_\mathfrak{p})$ for each $\mathfrak{p} \in S$, it follows that $x\mathfrak{o}_k, y\mathfrak{o}_k$ and $z\mathfrak{o}_k$ do not all have a common factor, so we will drop this condition from here on.

Hence it suffices to study

$$T(P) := \sum_{\substack{x\in\mathfrak{o}_k\\ |D^{-1}P^{-1}x-\zeta_0^{(1)}|_\tau<1\\ x\equiv a \bmod \mathfrak{m}\\ x\mathfrak{d}^{-1} \text{ prime}}} 1 = \sum_{\substack{\mathfrak{q} \text{ prime}\\ \mathfrak{q}\in[\mathfrak{d}]^{-1}\\ N(\mathfrak{q})\ll P^d}} \sum_{\substack{x\mathfrak{o}_k=\mathfrak{q}\mathfrak{d}\\ |D^{-1}P^{-1}x-\zeta_0^{(1)}|_\tau<1\\ x\equiv a \bmod \mathfrak{m}}} 1. \tag{7.3.4}$$

Over $\mathbb{Q}$ this can be estimated immediately from the prime number theorem for arithmetic progressions. However in the setting of a general number field the interplay between the congruence condition on x, the ideal conditions imposed on the prime $\mathfrak{q}$ and the lattice point constraint requires further work.

Lemma 7.3.6. *We have*

$$T(P) \gg \frac{P^d}{\log P}.$$

The strategy to prove this is to first estimate the number of prime ideals $\mathfrak{q}$ in the above sum for which the ideal $\mathfrak{q}\mathfrak{d}$ has a generator lying in the correct congruence class. Afterwards we will deduce that for such primes we can in fact find a generator of $\mathfrak{q}\mathfrak{d}$ satisfying all of the conditions of the inner sum above. During the course of the proof we will need to apply estimates for sums of Hecke characters.

Definition 7.3.7. A Hecke character is a continuous group homomorphism $\chi : \mathbb{A}_k^* \to S^1$ from the idele group to the unit disk such that $\chi(k^*) = 1$.

This is equivalent to the notion of a Größencharakter.

Definition 7.3.8. Let $\mathfrak{m}$ be an integral ideal in k, and denote by $J^\mathfrak{m}$ the set of integral ideals of k which are relatively prime to $\mathfrak{m}$. A Größencharakter (of conductor $\mathfrak{m}$) is a character on $J^\mathfrak{m}$ for which there exists a pair of characters $\chi_f : (\mathfrak{o}_k/\mathfrak{m})^\times \to S^1$ and $\chi_\infty : (k \otimes_\mathbb{Q} \mathbb{R})^\times \to S^1$ such that

$$\chi((a)) = \chi_f(a)\chi_\infty(a), \tag{7.3.5}$$

for every $a \in \mathfrak{o}_k$ which is relatively prime to $\mathfrak{m}$. We call χ_f the finite type and χ_∞ the infinity type of the Größencharakter.

We are now ready to begin the proof of Lemma 7.3.6.

Lemma 7.3.9. *Suppose $a \in \mathfrak{o}_k$ and $\mathfrak{m} \subset \mathfrak{o}_k$ are relatively prime. Let $\mathfrak{d} \subset \mathfrak{o}_k$. If we define*

$$\Sigma(X) := \#\{\mathfrak{q} \text{ prime} : N(\mathfrak{q}) \ll X, \mathfrak{q} \in [\mathfrak{d}]^{-1} \text{ and } \exists x \text{ s.t } x\mathfrak{o}_k = \mathfrak{q}\mathfrak{d} \text{ and } x \equiv a \bmod \mathfrak{m}\},$$

then we have

$$\Sigma(X) \gg_{k,\mathfrak{m}} \frac{X}{\log X}.$$

Proof. For each prime $\mathfrak{q}$ we make a (non-canonical) choice x of generator for the (necessarily principal) ideal $\mathfrak{q}\mathfrak{d}$. If this generator satisfies $x \equiv au \bmod \mathfrak{m}$ for some $u \in \mathfrak{o}_k^\times$ then there must exist a generator in the desired congruence class. Therefore

$$\Sigma(X) = \sum_{\substack{\mathfrak{q} \text{ prime} \\ N(\mathfrak{q}) \ll X \\ \mathfrak{q} \in [\mathfrak{d}]^{-1}}} \sum_{\substack{x \in \mathfrak{o}_k/\mathfrak{o}_k^\times \\ N(x) \leqslant X \\ x\mathfrak{o}_k = \mathfrak{q}\mathfrak{d}}} \mathbf{1}_{x \bmod \mathfrak{m} = au \text{ for } u \in \mathrm{Im}(\mathfrak{o}_k^\times \to \mathfrak{o}_k/\mathfrak{m})}.$$

Using orthogonality this indicator function can be expressed in terms of characters on $\mathfrak{o}_k^\times/\mathfrak{m}$. Namely

$$\mathbf{1}_{x \bmod \mathfrak{m} = au \text{ for } u \in \mathrm{Im}(\mathfrak{o}_k^\times \to \mathfrak{o}_k/\mathfrak{m})} = \frac{\#\mathrm{Im}(\mathfrak{o}_k^\times \to \mathfrak{o}_k/\mathfrak{m})}{\phi(\mathfrak{m})} \sum_{\substack{\chi_f \bmod \mathfrak{m} \\ \chi_f|_{\mathrm{Im}(\mathfrak{o}_k^\times \to \mathfrak{o}_k/\mathfrak{m})} \equiv 1}} \chi_f(x\overline{a}), \tag{7.3.6}$$

where $\overline{a}$ denotes the inverse of a modulo $\mathfrak{m}$. We approach these character sums by passing to Hecke characters. We will lift each of the characters χ_f appearing in (7.3.6) to a Hecke character of conductor $\mathfrak{m}$. Every Hecke character is a Größencharakter which must satisfy the expression

$$\chi_f(\epsilon)\chi_\infty(\epsilon) = 1 \quad \text{for each } \epsilon \in \mathfrak{o}_k^\times.$$

(Moreover any pair of characters satisfying this expression must come from a Größencharakter [88, Chapter VII, Exercise 5].) Therefore because the characters in (7.3.6) are trivial on units so must the infinity types χ_∞ be. Suppose that χ_1 and χ_2 are Hecke characters with finite type of the form χ_f in (7.3.6) and trivial infinity type. Then $\chi_1\chi_2^{-1}$ is a Hecke character of trivial infinity type and conductor 1. Hence (by for example [88, Chapter VII Corollary 6.10]) $\chi_1\chi_2^{-1}$ are characters of the class group. Therefore if one sums over all Hecke lifts of the χ_f with trivial infinity type, the result is 0 unless the argument is a principal ideal. Replacing the sum over finite characters in (7.3.6) by the sum over the induced Hecke characters yields

$$\Sigma(X) = \frac{1}{\phi(\mathfrak{m})h_k} \sum_{\substack{\mathfrak{q} \text{ prime} \\ N(\mathfrak{q}) \ll X}} \chi \sum_{\substack{\text{Hecke mod } \mathfrak{m} \\ \chi_\infty = 1}} \chi((\overline{a}))\chi(\mathfrak{q}\mathfrak{d}).$$

Finally we invoke the classical result (e.g. [87, Prop 7.17])

$$\sum_{\substack{\mathfrak{q} \text{ prime} \\ N(\mathfrak{q}) \leqslant X}} \chi(\mathfrak{q}) = \begin{cases} \frac{X}{\log X}\left(1 + o(1)\right) & \text{if } \chi \text{ principal} \\ o\left(\frac{X}{\log X}\right) & \text{otherwise.} \end{cases}$$

□

With this in hand we complete the proof of Lemma 7.3.6.

Proof of Lemma 7.3.6. Here as before we set $\mathfrak{d} = \gcd(a\mathfrak{o}_k, \mathfrak{m})$. We start with the case where the unit group of k is infinite. If so we can furnish a lower bound for the inner sum in (7.3.4) by fixing a generator x and counting the units u satisfying $|D^{-1}P^{-1}ux - \zeta_0^{(1)}|_\tau < 1$ and $u \equiv 1 \bmod \mathfrak{m}$. The group $U_k(\mathfrak{m})$ of units congruent to 1 $\bmod \mathfrak{m}$ is a congruence subgroup which is mapped to a lattice under the usual map $\mathfrak{o}_k^\times \to \mathbb{R}^{r+s}$. Since we are only interested in a lower bound we may restrict to just counting those x with $N(x) \leqslant P^{d/2}$. The inner sum in (7.3.4) then represents a lattice point count in an expanding region. For P large enough this is certainly $\geqslant 1$.

It now remains to count those primes in the outer sum for which a suitable generator x of $\mathfrak{q}\mathfrak{d}$ exists. The congruence in (7.3.4), however, need not be primitive so Lemma 7.3.9 does not directly apply yet. If $\mathfrak{q} \mid \mathfrak{d}$ is a prime ideal then x satisfying $x \equiv a \bmod \mathfrak{m}$ must lie in the ideal $\mathfrak{q}^{\mathrm{ord}_\mathfrak{q}(a)}$. Enlarging $\mathfrak{m}$ if necessary, we have that $\mathrm{ord}_\mathfrak{q}(\mathfrak{m}) > \mathrm{ord}_\mathfrak{q}(a)$. Let π be a uniformising element in $k_\mathfrak{q}$ then

$$a = \pi^{\mathrm{ord}_\mathfrak{q}(a)} u$$

where u is a $\mathfrak{q}$-adic unit. Since $x \in \mathfrak{q}^{\mathrm{ord}_\mathfrak{q}(a)}$ we can write $x = \pi^{\mathrm{ord}_\mathfrak{q}(a)} x'$ then

$$x \equiv a \bmod \mathfrak{q}^{\mathrm{ord}_\mathfrak{q}(\mathfrak{m})} \iff x' \equiv u \bmod \mathfrak{q}^{\mathrm{ord}_\mathfrak{q}(\mathfrak{m}) - \mathrm{ord}_\mathfrak{q}(a)}.$$

Let $\mathfrak{m}' = \mathfrak{m}/\mathfrak{d}$ then the above tells us that there exists a residue class $a' \in (\mathfrak{o}_k/\mathfrak{m}')^\times$ such that

$$\begin{aligned} T(P) &\geqslant \#\{\mathfrak{q} \text{ prime} : N(\mathfrak{q}) \ll P^d \text{ and } \exists x' \equiv a' \bmod \mathfrak{m}' \text{ s.t. } x'\mathfrak{o}_k = \mathfrak{q}\mathfrak{d}\} \\ &= \Sigma(P^d). \end{aligned}$$

Turning to the case of imaginary quadratic fields, we may write

$$T(P) \geqslant \frac{1}{\#\mathfrak{o}_k^\times} \sum_{\substack{\mathfrak{q} \text{ prime} \\ \mathfrak{q} \in [\mathfrak{d}]^{-1} \\ N(\mathfrak{p}) \ll P^d \\ (7.3.7)}} 1,$$

where the summands satisfy the conditions

$$\exists x \in \mathfrak{o}_k \text{ s.t. } |P^{-1}x - Dx_\infty|_\tau < 1, x \equiv a \bmod \mathfrak{m} \text{ and } x\mathfrak{o}_k = \mathfrak{q}\mathfrak{d}. \tag{7.3.7}$$

The condition that x lie in the box above is difficult to detect and so we will replace it by restricting the norm and argument of x instead. Since $|z|_\tau^2 \asymp N(z)$ for $z \in k$, there exists a constant C such that

$$N(x - PDx_\infty) < CP^2 \implies |P^{-1}x - Dx_\infty|_\tau < 1.$$

For any small parameter $\gamma > 0$, define α and β, respectively, to be the arguments of the two points of intersetion of the balls defined by the equations

$$\begin{cases} N(z) = CP^2 + N(PDx_\infty) - \gamma \\ N(z - PDx_\infty) = CP^2. \end{cases}$$

A sufficient condition for $N(x - PDx_\infty) < CP^2$ is that x lies in the region defined by

$$\begin{cases} N(PDx_\infty) - CP^2 + \gamma < N(z) < CP^2 + N(PDx_\infty) - \gamma \\ \alpha < \arg(z) < \beta. \end{cases}$$

To ease notation let $\mathscr{B}$ denote the interval $[N(PDx_\infty) - CP^2 + \gamma, N(PDx_\infty) + CP^2 - \gamma]$. We will detect the restriction on the argument using Hecke characters on k. In particular, define

$$\lambda^n(x) := \left(\frac{x}{N(x)}\right)^{n\#\mathfrak{o}_k^\times}.$$

Since this continuous homomorphism is trivial on units, it can be lifted to a Hecke character of conductor 1. By abuse of notation, we will fix such a Hecke character and refer to it as λ^n also. Let $I(x)$ be the indicator function that detects when $x \equiv a \bmod \mathfrak{m}$ and that the ideal generated by x is equal to $\mathfrak{q}\mathfrak{d}$ for some prime $\mathfrak{q}$. Similarly $\widetilde{I}$ will detect when a principal ideal has a generator satisfying these conditions. Writing $\mathbf{1}_{(\alpha,\beta)}(x)$ for the indicator of the interval, and setting $\lambda = \lambda^1$ we get the lower buond

$$\sum_{\substack{x \in \mathfrak{o}_k \\ N(x) \in \mathscr{B} \\ \alpha < \arg x < \beta}} I(x) \geqslant \frac{1}{\#\mathfrak{o}_k^\times} \sum_{\substack{\mathfrak{a} \subset \mathfrak{o}_k \text{ principal} \\ N(\mathfrak{a}) \in \mathscr{B}}} \widetilde{I(\mathfrak{a})} \mathbf{1}_{(\alpha,\beta)}(\arg\lambda(\mathfrak{a})).$$

This is because $\mathbf{1}_{(\alpha,\beta)}(\arg\lambda(\mathfrak{a}))$ will detect when the principal ideal $\mathfrak{a}$ has a generator which lies in the correct sector. Since $\mathbf{1}_{(\alpha,\beta)}$ is a $\frac{2\pi}{\beta-\alpha}$ periodic function it has a Fourier expansion

$$\mathbf{1}_{(\alpha,\beta)}(z) = \frac{\beta - \alpha}{2\pi} \sum_{n \in \mathbb{Z}} \widehat{\mathbf{1}_{(\alpha,\beta)}}(n) \exp(2\pi i n z),$$

where $\widehat{\mathbf{1}_{(\alpha,\beta)}}(n)$ dentes the n^{th} Fourier coefficient. Using the expresion $\exp(2\pi i \arg(z)) =$

$\frac{z}{N(z)}$, we can write

$$\sum_{\substack{x\in\mathfrak{o}_k\\ N(x)\in\mathscr{B}\\ \alpha<\arg x<\beta}} I(x) \geqslant \frac{\beta-\alpha}{2\pi\#\mathfrak{o}_k^\times} \sum_{\substack{\mathfrak{a}\subset\mathfrak{o}_k \text{ principal}\\ N(\mathfrak{a})\in\mathscr{B}}} \widetilde{I(\mathfrak{a})} \sum_{n\in\mathbb{Z}} \widehat{\mathbf{1}_{(\alpha,\beta)}}(n)\lambda^n(\mathfrak{a})$$

$$= \frac{\beta-\alpha}{2\pi\#\mathfrak{o}_k^\times}\left(\sum_{\substack{\mathfrak{a}\subset\mathfrak{o}_k \text{ principal}\\ N(\mathfrak{a})\in\mathscr{B}}} \widetilde{I(\mathfrak{a})} + \sum_{\substack{\mathfrak{a}\subset\mathfrak{o}_k \text{ principal}\\ N(\mathfrak{a})\in\mathscr{B}}} \widetilde{I(\mathfrak{a})} \sum_{n\neq 0} \widehat{\mathbf{1}_{(\alpha,\beta)}}(n)\lambda^n(\mathfrak{a})\right).$$

By definition the first sum is clearly equal to

$$\#\mathfrak{o}_k^\times \sum_{\substack{x\in\mathfrak{o}_k\\ N(x)\in\mathscr{B}}} I(x),$$

which we bound below by $\Sigma(P^2)$. To estimate the second term we replace $\chi_{(\alpha,\beta)}$ by an approximating function H which satisfies

- $\widehat{H}(n) = 0$ for $|n| > N$,
- $\int_{-\infty}^{\infty} \left(\chi_{(\alpha,\beta)} - H\right)(x)\mathrm{d}x = \frac{1}{N}$.

That such a function exists is guaranteed by [3, Lemma 2.5]. Hence, replacing $\widehat{\chi_{(\alpha,\beta)}}$ by $\widehat{H}$ we may bound the second sum by

$$\ll \frac{1}{N} \sum_{\substack{\mathfrak{a}\subset\mathfrak{o}_k \text{ principal}\\ N(\mathfrak{a})\in\mathscr{B}}} \widetilde{I(\mathfrak{a})} + \sum_{n=1}^{N} \left| \sum_{\substack{\mathfrak{a}\subset\mathfrak{o}_k \text{ principal}\\ N(\mathfrak{a})\in\mathscr{B}}} \widetilde{I(\mathfrak{a})}\lambda^n(\mathfrak{a}) \right|.$$

The first error term is negligible for any N which grows asymptotically as P does so we restrict attention to the second error term. Letting $c_{\mathfrak{a}}$ denote the indicator function for principal ideals with a generator satisfying the conditions covered by I and for having norm in $\mathscr{B}$, the Cauchy–Schwarz inequality gives

$$\left| \sum_{N(x)\leqslant P^2} c_{\mathfrak{a}}\lambda^n(x) \right| \leqslant \left| \sum_{N(\mathfrak{a})\leqslant P^2} c_{\mathfrak{a}}^2 \right|^{\frac{1}{2}} \left| \sum_{N(\mathfrak{a})\leqslant P^2} \lambda^{2n}(\mathfrak{a}) \right|^{\frac{1}{2}}.$$

Estimating the first sum trivially, the result follows by applying the Polya–Vinogradov inequality for Hecke characters (originally due to Landau [73]), namely

$$\left| \sum_{N(\mathfrak{a})\leqslant P^2} \lambda^{2n}(\mathfrak{a}) \right| \ll P^{2/3}.$$

□

We can now conclude that for sufficiently large P, $N(P) \gg \frac{P^{3d}}{\log^3 P}$ and thus the set $\mathscr{A}$ is non-empty as desired.

Bibliography

[1] M. Artin and D. Mumford, Some elementary examples of unirational varieties which are not rational. *Proc. London Math. Soc.* (3) **25** (1972), 75–95.

[2] A. Baily, On the density of discriminants of quartic fields. *J. reine angew. Math.* **315** (1980), 190–210.

[3] R. C. Baker, *Diophantine Inequalities*. London Mathematical Society Monographs. New Series, 1. Oxford Science Publications. The Clarendon Press, Oxford University Press, New York, 1986.

[4] V.V. Batyrev and Y. Tschinkel, Rational points on some Fano cubic bundles. *C. R.Acad. Sci. Paris* **323** (1996), 41–46.

[5] H.-J. Bartels, Zur Arithmetik von Konjugationsklassenin algebraischen Gruppen. *J. Algebra* **70** (1981), 179–199.

[6] H.-J. Bartels, Zur Arithmetik von Diedergruppenerweiterungen. *Math. Ann.* **256** (1981), 465–473.

[7] A. Beauville, J.-L. Colliot-Thélène, J.-J. Sansuc and P. Swinnerton-Dyer, Variétés stablement rationnelles nonrationnelles. *Ann. of Math.* **121** (1985), 283–318.

[8] M. Bhargava, A positive proportion of plane cubics fail the Hasse principle. Pre-print, arXiv:1402.1131.

[9] M. Bhargava, J. Cremona, and T. Fisher, The proportion of plane cubic curves over $\mathbb{Q}$ that everywhere locally have a point. *Int. J. Number Theory* **12** (2016), no. 4, 1077–1092.

[10] Bryan J. Birch, Forms in many variables. *Proc. Roy. Soc. Ser. A* **265** (1962), 245-263.

[11] P. Le Boudec, T. D. Browning and W. Sawin, The Hasse principle for random Fano hypersurfaces. Pre-print, arXiv:2006.02356.

[12] R. la Bretèche, P. Kurlberg, and I. Schparlinski, On the number of products which form perfect powers and discriminants of multiquadratic extensions. *IMRN*, to appear. arXiv:1901.10694.

[13] R. la Bretèche and T. D. Browning, Density of Châtelet surfaces failing the Hasse principle. *Proc. Lond. Math. Soc.* (3) **108** (2014), no. 4, 1030–1078.

[14] M. Bright, Rational points in families of varieties. Lecture notes from *Autumn School on Algebraic and Arithmetic Geometry*, Johannes Gutenberg-Universität Mainz.

[15] M. Bright, T. D. Browning and D. Loughran, Failures of weak approximation in families. *Compositio Math.* **152** (2016), 1435–1475.

[16] T. D. Browning and A. Gorodnik, Power-free values of polynomials on symmetric varieties, *Proc. Lond. Math. Soc.* **114** (2017) 1044–1080.

[17] Density of rational points on a quadric bundle in $\mathbb{P}^3 \times \mathbb{P}^3$. Preprint, arXiv:1805.10715.

[18] T. D. Browning and D. R. Heath-Brown, Quadratic polynomials represented by norm forms. *Geom. Funct. Anal.* **22** (2012), no. 5, 1124–1190.

[19] T. D. Browning and D. Loughran, Sieving rational points on varieties. *Trans. Amer. Math. Soc.* **371** (2019), no. 8, 5757–5785.

[20] T. D. Browning and L. Matthiesen, Norm forms for arbitrary number fields as products of linear polynomials. *Ann. Sci. Éc. Norm. Supér.* (4) **50** (2017), no. 6, 1383–1446.

[21] T. D. Browning and R. Newton, The proportion of failures of the Hasse norm principle. *Mathematika* **62** (2016), no. 2, 337–347.

[22] T. D. Browning and W. Sawin, Free rational curves on low degree hypersurfaces and the circle method. Pre-print, arXiv:1810.06882.

[23] T. D. Browning and P. Vishe, Rational curves on smooth hypersurfaces of low degree. *Algebra Number Theory* **11**(7) (2017), 1657–1675.

[24] J.W.S Cassels and A. Frölich, *Algebraic Number Theory.* Academic Press, 1967.

[25] C. Clemens and P. Griffiths, The intermediate Jacobian of the cubic threefold. *Ann. of Math.* (2) **95** (1972), 281–356.

[26] H. Cohen, F. Diaz y Diaz, and M. Olivier, A survey of discriminant counting. *Algorithmic Number Theory*, pp 80–94, Springer, 2002.

[27] S. D. Cohen, The distribution of Galois groups and Hilbert's irreducibility theorem. *Proc. London Math. Soc.* **3**(1981), 227–250.

[28] H. Cohn, *Advanced Number Theory*, Dover Publications, 1980.

[29] J.-L. Colliot-Thélène, Points rationnels sur les fibrations. *Higher dimensional varieties and rational points (Budapest, 2001)*, Bolyai Soc. Math. Stud., vol. 12, Springer, Berlin, 2003, pp. 171–221.

[30] J.-L. Colliot-Thélène and A. Pirutka, Hypersurfaces quartiques de dimension 3: non-rationalité stable. *Ann. Sci. Éc. Norm. Supér.* (4) **49** (2016), no. 2, 371–397.

[31] J.-L. Colliot-Thélène, D. Coray and J.-J. Sansuc, Descent and the Hasse principle for certain rational varieties, *J. reine angew. Math.* **320** (1980) 150–191.

[32] J.-L. Colliot-Thélène, D. Harari, and A. N. Skorobogatov, Valeurs d'un polynôme à une variable représentées par une norm. *Number Theory and Algebraic Geometry.* London Math. Soc. Lecture Note Ser., vol. 303, pp. 69–89. Cambridge Univ. Press, Cambridge (2003).

[33] J.-L. Colliot-Thélène and P. Salberger, Arithmetic on some singular cubic hypersurfaces. *Proc. London Math. Soc.* (3) **58** (1989), no. 3, 519–549.

[34] J.-L. Colliot-Thélène and J.-J. Sansuc, La descente sur les variétés rationnelles. Journées de Géometrie Algébrique d'Angers, Juillet 1979 pp. 223–237, Sijthoff & Noordhoff, Alphen aan den Rijn—Germantown, Md., 1980.

[35] J.-L. Colliot-Thélène, J.-J. Sansuc, and P. Swinnerton-Dyer, Intersections of two quadrics and Châtelet surfaces I. *J. reine angew. Math.* **373** (1987), 37–107.

[36] J.-L. Colliot-Thélène, J.-J. Sansuc, and P. Swinnerton-Dyer, Intersections of two quadrics and Châtelet surfaces II. *J. reine angew. Math.* **374** (1987), 72–168.

[37] J.-L. Colliot-Thélène and A. N. Skorobogatov, *The Brauer–Grothendieck group.* To be published, available at `http://wwwf.imperial.ac.uk/~anskor/brauer.pdf`.

[38] J.-L. Colliot-Thélène, A. N. Skorobogatov, and P. Swinnerton-Dyer, Double fibres and double covers: paucity of rational points. *Acta Arith.* **79** (1997), no. 2, 113–135.

[39] J.-L. Colliot-Thélène, and P. Swinnerton-Dyer, Hasse principle and weak approximation for pencils of Severi-Brauer and similar varieties. *J. Reine Angew. Math.* **453** (1994), 49–112.

[40] J.L. Colliot-Thélène and F. Xu, Strong approximation for the total space of certain quadratic fibrations. *Acta Arith.* **157** (2013), no. 2, 169–199.

[41] H. Davenport, *Multiplicative number theory. Third edition.* Graduate Texts in Mathematics, Springer-Verlag, New York, **74** (2000).

[42] U. Derenthal, A. Smeets, and D. Wei, Universal torsors and values of quadratic polynomials represented by norms. *Math. Ann.* **361** (2015), no. 3-4, 1021–1042.

[43] J. Franke, Y.I. Manin and Y. Tschinkel, Rational points of bounded height on Fano varieties. *Invent. Math.* **95** (1989), 421–435.

[44] C. Frei, D. Loughran, and R. Newton, The Hasse norm principle for abelian extensions. *Amer. J. Math.* **140**(6) (2018) 1639–1685.

[45] C. Frei, D. Loughran, and R. Newton, Number fields with prescribed norms. Pre-print, arXiv:1810.06024.

[46] J. Friedlander and H. Iwaniec, *Opera de Cribro*, Amer. Math. Soc., 2004.

[47] J. Friedlander and H. Iwaniec, Ternary quadratic forms with rational zeros. *Journal de Théorie des Nombres de Bordeaux* **22** (2010) 97–113.

[48] R. Fritsch, Counting multi-quadratic number fields of bounded discriminant. Preprint, arXiv:1902.03202.

[49] A. Frölich and M. J. Taylor, *Algebraic Number Theory.* Cambridge University Press, 1991.

[50] O. Gabber, Some theorems on Azumaya algebras. *The Brauer group (Sem., Les Plans-sur-Bex, 1980)*, pp. 129–209, Lecture Notes in Math., 844, Springer, Berlin-New York, 1981.

[51] P. Gille and T. Szamuely, *Central simple algebras and Galois cohomology.* Cambridge Studies in Advanced Mathematics, **165**. Cambridge University Press, Cambridge, 2017.

[52] L. Goldmakher and B. Louvel, A quadratic large sieve inequality over number fields. *Math. Proc. Camb. Phil. Soc.*, **154**(2) (2013), 193–212.

[53] C. R. Guo, On solvability of ternary quadratic forms. *Proc. London Math. Soc.* (3) **70** (1995) 241–263.

[54] D. Harari, Méthode des fibrations et obstruction de Manin. *Duke Math. J.* **75** (1994), 221–260.

[55] R. Hartshorne, *Algebraic geometry.* Graduate Texts in Mathematics, Springer-Verlag, New York-Heidelberg, **52**, (1977).

[56] G. Harman and P. Lewis, Gaussian primes in narrow sectors. *Mathematika*, **48**(1-2) (2001), 119–135.

[57] Y. Harpaz and O. Wittenberg, On the fibration method for zero-cycles and rational points. *Ann. of Math.* (2) **183** (2016), no. 1, 229–295.

[58] H. Hasse, Beweis eines Satzes und Wiederlegung einer Vermutung über das allgemeine Normenrestsymbol. *Nachr. Ges. Wiss. Göttingen Math.-Phys. Kl.*(1931) 64–69.

[59] B. Hassett, A. Pirutka and Y. Tschinkel, Stable rationality of quadric surface bundles over surfaces. *Acta Math.* **220** (2018), no. 2, 341–365.

[60] W. Heap, M. Radziwill and K. Soundararajan, Sharp upper bounds for fractional moments of the Riemann zeta function. To appear *Q. J. Math.* (2019), arXiv:1901.08423.

[61] D. R. Heath-Brown, A mean value estimate for real character sums. *Acta Arith.***72** (1995), no. 3, 235–275.

[62] D. R. Heath-Brown and A. N. Skorobogatov, Rational solutions of certain equations involving norms. *Acta Math.* **189** (2002), no. 2, 161–177.

[63] D. Holmes, N. Rome, Fields of definition of curves of a given degree. *Preprint.* arXiv:1901.11294.

[64] C. Hooley, On ternary quadratic forms that represent zero I. *Glasg. Math. J.* **35** (1993) 13–23.

[65] C. Hooley, On ternary quadratic forms that represent zero II. *J. reine angew. Math.* **602** (2007) 179–225.

[66] V. A. Iskovskikh, A counterexample to the Hasse principle for systems of two quadratic forms in fi

ve variables. *Mat. Zametki* **10** (1971) 253–257.

[67] V. A. Iskovskikh and Y. I. Manin, Three-dimensional quartics and counterexamples to the Lüroth problem. *Math. USSR-Sb.***15** (1971), 141–166.

[68] H. Iwaniec and E. Kowalski, *Analytic Number Theory.* AMS Colloquium Publ. 53, 2004.

[69] J. Kollár, Unirationality of cubic hypersurfaces. *J. Inst. Math. Jussieu* **1** (2002), no. 3, 467–476.

[70] M. Kontsevich and Y.I. Manin, Gromov-Witten classes, quantum cohomology, and enumerative geometry. *Comm. Math. Phys.*, **164**(3) (1994), 525–562.

[71] A. Lachand, *Entiers friables et formes binaires.* Ph.D thesis, Universitéde Lorraine, 2014.

[72] J. Lagarias and K. Soundararajan, Counting smooth solutions to the equation $A + B = C$. *Proc. Lond. Math. Soc.* (3) **104** (2012), no. 4, 770–798.

[73] E. Landau, Über Ideale und Primideale in Idealklassen. *Math. Z.* **2** (1918), no. 1-2, 52–154.

[74] C. Le Rudulier, *Points algébriques de hauteur bornée.* Ph.D thesis, Université de Rennes 1, 2014.

[75] C.-E. Lind, Untersuchungen über die rationalen Punkte der ebenen kubischen Kurven vom Geschlecht Eins, Diss. Univ. Uppsala 1940.

[76] D. Loughran, The number of varieties in a family which contain a rational point. *J. Eur. Math. Soc.* **20** (2018), no. 10, 2539–2588.

[77] D. Loughran and L. Matthiesen, Frobenian multiplicative functions and rational points in fibrations. Preprint, arXiv:1904.12845.

[78] D. Loughran and A. Smeets, Fibrations with few rational points. *Geom. Funct. Anal.* **26** (2016), no. 5, 1449–1482.

[79] D. Loughran, R. Takloo-Bighash, and S. Tanimoto, Zero-loci of Brauer group elements on semi-simple algebraic groups. *Journal de l'Institut de Mathématiques de Jussieu*, to appear, arXiv:1705.09244.

[80] J. Lüroth, Beweis eines Satzes über rationale Curven. *Math. Ann.* **9** (1876), 163–165.

[81] A. Macedo, The Hasse norm principle for A_n-extensions. *J. Num. Theo.* (2019).

[82] A. Macedo and R. Newton, Explicit methods for the Hasse norm principle and applications to A_n and S_n extensions. Preprint, arXiv:1906.03730.

[83] S. Mäki, The conductor density of abelian number fields. *J. Lond. Math. Soc.* **47** (1993), no. 2, 18–30.

[84] Y.I. Manin, Le groupe de Brauer–Grothendieck en géométrie diophantienne. *Actes Congrès Int. Math. Nice 1970*, tome **I**, Gauthier-Villars, Paris 401–411 (1971).

[85] M. Matchett-Wood, On the probabilities of local behaviors in abelian field extensions. *Comp. Math.* **146** (2010), no. 1, 102–128.

[86] Y. Matiyasevich, The Diophantineness of enumerable sets. *Dokl. Akad. Nauk SSSR* **191** (1970) 279–282.

[87] W. Narkiewicz, *Elementary and analytic theory of algebraic numbers.* Springer Monographs in Mathematics, Springer-Verlag, Berlin-Heidelburg, (2004).

[88] J. Neukirch, *Algebraic number theory.* Translated from the 1992 German original and with a note by Norbert Schappacher. With a foreword by G. Harder. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], 322. Springer-Verlag, Berlin, (1999).

[89] E. Peyre, Hauteurs et mesures de Tamagawa sur les variétiés de Fano. *Duke Math. J.* **79** (1995), 101–218.

[90] M. Pieropan, A. Smeets, S. Tanimoto, A. Várilly-Alvarado, Campana points of bounded height on vector group compactifications. Preprint, arXiv:1908.10263.

[91] A. Pirutka, Algebraic geometry: Salt Lake City 2015, 459–483, Proc. Sympos. Pure Math., 97.2, Amer. Math. Soc., Providence, RI, 2018.

[92] V. Platanov and A. Rapinchuk, *Algebraic Groups and Number Theory.* Academic Press, 1993.

[93] B. Poonen, *Rational points on varieties.* Graduate Studies in Mathematics 186, American Mathematical Society, Providence, RI, 2017.

[94] B. Poonen and F. Voloch, Random Diophantine equations. With appendices by Jean-Louis Colliot-Thélène and Nicholas M. Katz. Progr. Math., **226**, *Arithmetic of higher-dimensional algebraic varieties (Palo Alto, CA, 2002)*, 175–184, Birkhäuser Boston, Boston, MA, 2004.

[95] Z. Ran, Enumerative geometry of singular plane curves. *Invent. Math.*, **97**(3) (1989), 447–465.

[96] H. Reichardt, Einige im Kleinen überall lösbare, im Grossen unlösbare diophantische Gleichungen. *J. Reine Angew. Math.* **184** (1942), 12–18.

[97] E. Robson, Words and pictures: new light on Plimpton 322. *Amer. Math. Monthly* **109** (2002), no. 2, 105–120.

[98] N. Rome, A positive proportion of Hasse principle failures in a family of Châtelet surfaces. *International Journal of Number Theory.* **15** (2019), no. 6, 1237–1249.

[99] N. Rome, The Hasse Norm Principle in Biquadratics. *Journal de Théorie des Nombres de Bordeaux.* **30** (2018), no. 3, 947–964.

[100] D. Schindler, Manin's conjecture for certain biprojective hypersurfaces. *J. Reine Angew. Math.* **714** (2016), 209–250.

[101] D. Schindler and A. N. Skorobogatov, Norms as products of linear polynomials. *J. London Math. Soc.* **89** (2014), 559–580.

[102] Damaris Schindler and Efthymios Sofos, Sarnak's saturation problem for complete intersections. *Mathematika* **65** (2019), no. 1, 1–56.

[103] S. Schreieder, On the rationality problem for quadric bundles. *Duke Math. J.* **168** (2019), no. 2, 187–223.

[104] E. Selmer, The Diophantine equation $ax^3 + by^3 + cz^3 = 0$. *Acta Mathematica* **85** (1951), 203–362.

[105] J.-P. Serre, *A course in arithmetic.* Graduate Texts in Mathematics, Springer-Verlag, New York-Heidelburg, **7**, (1973).

[106] J.-P. Serre, *Lectures on the Mordell-Weil theorem. Third edition.* Aspects of Mathematics, Friedr. Vieweg & Sohn, Braunschweig, (1997).

[107] J.-P. Serre, *Local fields*. Graduate Texts in Mathematics, Springer-Verlag, New York-Heidelburg, **67**, (1979).

[108] J.-P. Serre, Spécialisation des éléments de $\mathrm{Br}_2(\mathbb{Q}(T_1,\dots,T_n))$. *C. R. Acad. Sci. Paris Sér. I Math.* **311** (1990) 397–402.

[109] J.-P. Serre, *Topics in Galois theory*. AK Peters/CRC Press, (2016).

[110] C. Skinner, Forms over number fields and weak approximation. *Compos. Math.*, **102**(1) (1997), 11–29.

[111] A. N. Skorobogatov, Beyond the Manin obstruction. *Invent. Math.* **135** (1999), no. 2, 399–424.

[112] A.N. Skorobogatov, Arithmetic on certain quadric bundles of relative dimension 2. I. *J. reine angew. Math.* **407** (1990), 57–74.

[113] E. Sofos, Serre's problem on the density of isotropic fibres in conic bundles. *Proc. Lond. Math. Soc.* (3) **113** (2016), no. 2, 261–288.

[114] E. Sofos and E. Visse, The density of fibres with a rational point for a fibration over hypersurfaces of low degree. To appear, *Annales de l'Institut Fourier*, arXiv:1804.05768.

[115] M. Swarbrick-Jones, A note on a theorem of Heath-Brown and Skorobogatov. *Q. J. Math.* **64** (2013), no. 4, 1239–1251.

[116] H.P.F. Swinnerton-Dyer, Two special cubic surfaces. *Mathematika* **9** (1962), 54–56.

[117] G. Tenenbaum, *Introduction to Analytic and Probabilistic Number Theory.* Cambirdge University Press, 1995.

[118] E. C. Titchmarsh, *The theory of the Riemann zeta-function. Second edition. Edited and with a preface by D. R. Heath-Brown.* The Clarendon Press, Oxford University Press, New York, 1986.

[119] I. Vainsencher, Enumeration of n-fold tangent hyperplanes to a surface. *J. Alg. Geom.* **4** (1995), 503–526.

[120] C. Voisin, Unirational threefolds with no universal codimension 2 cycle. *Invent. Math.* **201** (2015), no. 1, 207–237.

[121] V. Voskresenskiĭ, Birational properties of linear algebraic groups. *Izv. Akad. Nauk SSSR Ser. Mat.* **34** (1970,) 3–19. English translation: *Math. USSR-Izv.* Vol.4 (1970), 1–17.

[122] V. Voskresenskiĭ, B. Kunyavskiĭ, Maximal tori in semisimple algebraic groups. Manuscript deposited at VINITI 15.03.84, no.1269–84, 28pp.

[123] E. Wirsing, Das asymptotische Verhalten von Summen über multiplikative Funktionen. II *Acta Math. Acad. Sci. Hungar.* **18** (1967), 411–467.

[124] D. J. Wright, Distribution of discriminants of abelian extensions. *Proc. Lond. Math. Soc.* **58** (1989), no. 1, 17–50.

[125] HG Zeuthen. Almindelige egenskaber ved systemer af plane kurver, kongelige danske videnskabernes selskabs skrifter—naturvidenskabelig og mathematisk, 10 (1873) 285–393.